Science Library

MAY 11 1948

Vol. 38, No. 2 April, 1948

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Bryn Mawr College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

Response Fixation under Anxiety and Non-Anxiety Conditions: I. E. FARBER 111

An Investigation into the Causes of Retroactive Inhibition: C. E. OSGOOD ... 132

A Probability Formulation of the Hunter-Sigler Effect: H. SCHLOSBERG 155

'Superstition' in the Pigeon: B. F. SKINNER 168

Discriminative Conditioning. II. Effects of a Pavlovian Conditioned Stimulus upon a Subsequently Established Operant Response: W. K. ESTES .. 173

The Effects of Mild Anoxia on Simple Psychomotor and Mental Skills: R. W. RUSSELL 178

Visual Field Articulation in the Absence of Spatial Stimulus Gradients: C. R. BROWN AND J. W. GEBHARD 188

Discussion: Concerning the Effect of Shock for Right Responses in Visual Discrimination Learning: K. F. MUENZINGER, 201. *A Reply to Dr. Muenzinger on the Effect of Punishment on Discrimination Learning in a Non-Correction Situation:* G. J. WISCHNER, 203.

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in the Act of February 28, 1925, embodied in paragraph 4, Section 538, P. L. & R., authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.25), single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints, without covers, in the case of articles of seven or more printed pages, and with five free copies of the issue in the case of shorter contributions. One-half of the cost of tables, cuts, formulas, and author's alterations in galley proof is charged to the author.

Immediate publication can be given to articles whose authors assume the full cost of publication.

Address articles submitted for publication to the editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. ANDERSON and W. L. VALENTINE, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, **41**, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *E*s, *E*'s, and *E*s'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts should be typed double-spaced throughout, *including tables, footnotes, and bibliography*. Tables, figures, and footnotes should be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 38, NO. 2 APRIL, 1948

RESPONSE FIXATION UNDER ANXIETY AND NON-ANXIETY CONDITIONS *

BY I. E. FARBER

From the Psychological Laboratory, University of Iowa

INTRODUCTION

A number of experimental studies with rats have revealed that responses which occur in close temporal contiguity with shock may acquire unusual strength, tending to persist with little variability even though they do not lead to a goal and are therefore no longer adaptive. Such persistent non-adaptive behavior has been termed 'fixation,' and has been found to occur under a variety of conditions.

Among the most frequently investigated of these have been shock situations designed to produce in rats a phenomenon comparable to that of regression in human *Ss*. The relationship of fixation to regression has been indicated by Hamilton and Krechevsky, who define regression as "a reversion to an earlier, well-established mode of behavior, and persistence in that mode despite the relative inefficiency of that behavior in solving the problem . . ." (**3**, p. 238). Thus defined, regression becomes merely a special case of fixation. In the typical regression experiment an animal is trained to select one of two or more alternative approaches to a goal; this habit is then weakened by failure or by delay of reward and a second habit is established. Shock is then introduced into the situation; if the animal reverts to its first habit, it is said to have regressed.

The experimental investigations concerned with this problem have not shown unequivocally that regression occurs under these conditions (**1**, **3**, **6**, **10**, **14**, **18**, **20**). The weight of the evidence indicates, rather, that the

* This article is based on a dissertation submitted to the faculty of the Department of Psychology of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The writer is greatly indebted to Professor Kenneth W. Spence for his advice and assistance during the course of this study.

occurrence of regression is a direct function of the relative strength of previously established habits and an inverse function of the relative strength of on-going responses. The results with reference to fixation, however, have been consistent. Whatever its effect upon the direction of response, shock tends to constrict behavior. The initial response or responses elicited by shock serve as the behavior which is fixated, whether it be a previously made response, the ongoing response, or a response never before made by the animal in that situation.

The major findings with respect to the variables influencing fixation have been reviewed by Everall (1), Kleemeier (6), Mowrer (14), and Sears (21). The following discussion is therefore concerned only with a brief analysis of certain methodological considerations in the study of fixation as a function of shock.

Much of the evidence for fixation derives from situations in which an animal shows little variability of performance *during shock trials*, while taking the spatially or temporally longer of two or more possible paths to a goal (3, 6, 10, 20).[1] Thus, Hamilton and Krechevsky (3) found, during 100 trials in a simple maze involving a long and a short path to food, that animals subjected to shock showed only about one-ninth as much variability of performance as non-shock animals. Increased resistance to extinction of a response maintained under shock conditions has also been demonstrated. In Everall's experiment (1), when a previously established maze habit was no longer rewarded, animals shocked during the extinction period shifted to the correct response far more slowly than did non-shock control animals.

Kleemeier (6) has reported similar results using a four-choice maze. After successively establishing two different habits, both affording immediate access to food, 16 out of 20 of his animals persisted in their second habit during 30 or more shock trials (or reverted to it after first selecting another alley pointed in the same general direction). When a delay was introduced in the selected alley during an additional 20 to 30 shock trials, 96 percent of the runs were perseverative. When the locus of the shock was shifted to a point after choice during yet another 20 to 30 trials, only one animal showed any decided shift in preference, 93 percent of the group's runs continuing as before.

As Kleemeier has observed, these results demonstrate only that "it is more difficult to effect a change in preference while the animal is being shocked . . . than when the animal is not given such a shock" (6, p. 5); he prefers, therefore, to reserve the term fixation for persistence in a response "when shock [is] no longer given and the demands of the new situation [are] such that the mode of response adopted to shock [is] made inadequate" (6, p. 5). O'Kelly (18), using rats in an open-field situation, has demonstrated this sort of fixation; five of seven of his shock animals persisted in a previously established response for 16 to 29 trials after shock was removed. His control animals required an average of only 6.8 trials to

[1] Perhaps Everall's study (1) should also be included among these. Although she gave no reward in the longer alley, her discussion does not clearly indicate whether a correction procedure was used. If it was, the longer path involved merely a delay of reward rather than an absence of reward.

shift to the correct response. The two shock animals which learned the new response required 12 and 15 trials respectively. Kleemeier (6), in a second experiment, established two successive responses in 60 rats in his four-choice maze, and then gave 50 shock trials. The mean number of runs of these animals to the particular alley selected by each during the shock trials was 44.5, a score identical with that made by a control group of 10 non-shock animals. However, when shock was then removed and a delay of reward introduced for the ongoing response, the group which had been shocked averaged 29 perseverative responses whereas the non-shock group averaged only five. Nine of the shock animals persisted in their fixated response during the entire 50 trials.

To summarize, only two studies (6, 18) indicate anything more about the nature or degree of fixation as a function of shock than the fact that animals typically do not quickly learn to discriminate between long and short paths to food while they are being shocked. Of the studies investigating persistence of response after the removal of shock, only one (6) indicates the number of trials required for the extinction of the fixated response in a majority of animals; and in this instance retraining was discontinued after 50 trials.

Despite some lack of uniformity in the application of criteria to a definition of fixation, the experimental results clearly indicate that behavior under shock conditions may become highly rigid and resistant to extinction, even though alternative responses are made to have greater value in terms of reward or length of delay. This kind of behavior has been subjected to two sharply conflicting interpretations. One interpretation considers it to be a learning phenomenon, whereas the other regards it as abnormal—qualitatively different from learned behavior and explicable only in terms quite independent of those applied to learning phenomena.

The former interpretation of fixation has been best formulated by Sears (21, pp. 76 ff.), who identifies it with the concept of habit strength. Noting that experimental investigations have typically been concerned with instrumental act fixation rather than with the object or drive fixations so widely dealt with by Freud, he observes:

> As habit strength is customarily measured, it is equivalent to strength of instrumental act, and this latter is the quantitative continuum at one end of which lies what Freud called fixation, i.e., great strength.
>
> The chief measure of instrumental act strength is *degree of resistance to change* either by extinction (nonreward) or by retraining (reward of an incompatible habit). This is essentially the criterion of fixation used in the clinical situation . . . (21, p. 81).

Sears has listed a number of different factors which have been shown to influence the strength of instrumental acts in white rats; in his summary, punishment by means of shock is included among such other factors as amount of reinforcement and interval between instrumental act and goal response.

That fixation is a phenomenon which may be subsumed under the concept of habit strength has also been suggested by Mowrer, who has observed that in humans the habits that are acquired at successive stages of development necessarily have certain strengths, "and it is the varying strength of such habits that determines the extent to which *relative* fixation may be said to have occurred at any given point" (**14**, p. 58).

Certain of the experimental evidence strongly favors the notion that fixations which occur under shock conditions are merely relative and differ only in degree of strength from other habits. Animals subjected to shock do not by any means fall into two distinct groups—those which fixate and those which do not. The curves indicating degree of resistance to extinction of a response in shock animals are continuous, and in fact, typically overlap those for animals not subjected to punishment. Everall (**1**), for instance, found that shock animals had fewer hypotheses and maintained them for a longer period than did non-shock animals, but nevertheless observed that there was great variability in this respect, some animals tending to fixate regardless of the conditions under which they were run.

In sharp opposition to this point of view is that of Maier and his colleagues (**5**, **6**, **7**, **8**, **9**). Maier (**7**) has suggested that what is conventionally called the 'learning function' actually often involves processes other than learning. Such extraneous processes, including the "tendency to persist in an acquired mode of behavior after it ceases to be adaptive" (**7**, p. 250), should properly, according to his view, be studied as a problem independent of learning.[2]

Martin has differentiated fixation from learning on the grounds that fixation or "habit stability" is "determined primarily by factors of frequency, recency, and intensity," whereas "learning . . . seems to be primarily purposive and adjustive . . ." (**10**, p. 13). It is somewhat surprising to find here the suggestion that a phenomenon be distinguished from learning because it appears to be influenced by the very factors which are regarded by most learning theorists as determinants of habit strength. However, Tolman (**22**, pp. 59 ff.), has urged a distinction between learning and the 'psychological dynamisms' on somewhat similar grounds. He states:

> Learning is . . . a 'reasonable' activity which tends to keep the individual well adjusted to the actual environmental realities. . . . The dynamisms are, however, far less 'reasonable' than learning. . . . Learning corrects itself when the environmental relationships change. The dynamisms, on the other hand, tend to persist in hardened and blind form (pp. 60–61).
>
> We shall understand by fixation the tendency, after an auxiliary response or technique has been learned (because of its instrumental value), for such response or technique to

[2] Much of the discussion relevant to these differing interpretations has been concerned with the fixations which occur in animals in situations involving insoluble problems rather than noxious stimulation. However, proponents of the view that fixation involves a mechanism other than learning have explicitly included both under the rubric of abnormal phenomena (**5**, **6**, **8**). It is interesting to note in this connection the frequent use of air-blast in the insoluble problem situation. When Klee (**5**) eliminated air-blast as a factor under one experimental condition, he obtained fixation in only two out of 12 animals in an insoluble problem situation, whereas with air-blast nine out of 12 animals fixated. It is not suggested, however, that resistance to extinction in such situations is entirely a function of punishment; data from other sources (**17**) indicate that the pattern of reinforcement may be an important factor.

become relatively fixed and permanent—even though conditions may change so that it no longer has this original auxiliary value. This is the type of process for which Allport has coined the term 'functional autonomy' (pp. 61–62).

In opposition to this view McClelland has pointed out the gratuitousness of the assumption that certain acts must be performed for their own sake if they continue for an extended period after the conditions which originally instigated or maintained them have apparently been removed. As he puts it, "If it can be shown that there are good reasons why the act should continue, there will be no need to postulate a force within the act that keeps it going" (**11**, p. 275).

Following out this general line of thought, it might be assumed, for instance, that some form of secondary motivation and secondary reward could operate to produce such behavior phenomena. Thus, in the case of shock-produced fixation it would not seem unreasonable to suppose that the maintenance of the response is a consequence of secondary reinforcement due to the reduction of anxiety which has been conditioned to the stimulus cues at the locus of the shock.

This conception of anxiety reduction as a secondary reinforcing state has been elaborated in a number of papers by Mowrer (**13**, **15**, **16**) and by Miller and Dollard (**12**, p. 57 ff.). It assumes, in brief, that intense stimuli such as shock evoke in the organism certain responses, mediated largely by the sympathetic division of the autonomic nervous system, which produce an internal drive state. External cues (as well as internal cues) occurring at the time of shock acquire the capacity to evoke these visceral responses and thereby the drive, which, when thus elicited, is termed 'anxiety.'

A response which removes these conditioned cues would reduce the anxiety aroused by them and so strengthen the connection between these cues and any response which occurs in temporal contiguity with them, in accordance with the principle of reinforcement. According to this schema, the occurrence of the goal response (escape from pain or anxiety) strengthens first of all the tendency for the drive state to elicit the responses leading to the drive reduction. It also strengthens the connections between the cues present at the time of shock and the responses producing the drive state, between the cues and the escape response, and under certain conditions,[3] between the cues and the response actually leading to the noxious stimulus.

A hungry animal which escapes shock by entering a goal box in which it also obtains food is in effect being doubly reinforced. If the food is removed but the shock continued, there is no reason why, apart from the possible disruptive effect of the changed goal situation, the response leading to escape should not continue. Indeed, its strength should continue to

[3] Ordinarily, any increment of strength in the response to approach a noxious stimulus will not suffice to overcome the much greater increment in strength of the response which escapes it; when these two responses are incompatible, the escape response will almost invariably dominate. However, under at least two circumstances the approach responses may actually gain in relative strength: (a) if the situation is such that punishment is unavoidable, so that the only possible mode of escape is by way of an initial approach response; or (b) if the approach response already has considerable strength at the time punishment is introduced, so that it actually occurs a number of times, followed by escape from the punishment. Under these conditions, the complete sequence CUE → APPROACH → ANXIETY → ESCAPE may become very strong.

increase, though perhaps by smaller increments. More notably, even if the shock also is discontinued, the response would still receive secondary reinforcement due to the reduction of anxiety following escape from the cues associated with the shock. It is apparent then that if the mechanisms of secondary motivation and reinforcement are operative, the extinction of responses as a result of failure of food reinforcement should by no means proceed as rapidly in shock animals as in non-shock animals.

Statement of Problem

Among the implications of the theory that the fixation of non-adaptive responses in a shock situation may be the result of uncontrolled secondary reinforcement resulting from anxiety reduction is the following: if the anxiety were eliminated, by one means or another, the possibility of the maintenance of such responses by secondary reinforcement would be precluded. For instance, feeding in a situation in which anxiety has developed might remove the anxiety state, thereby eliminating the operation of secondary reinforcement and so decreasing the resistance to extinction of any non-adaptive response.

The present study compared the resistance to extinction of responses in a single-unit T-maze in two groups which had been subjected to shock, one of which had then been fed at the locus of shock, and the second of which had not. As a control for the possibility that the feeding activity competed with the responses directly rather than by way of reducing anxiety, a similar comparison was made between the resistance to extinction in two groups of animals not subjected to shock, but otherwise trained under conditions identical with those for the two shock groups.

The rationale of this study may be summarized as follows: when, under shock conditions, fixation of a response could be expected to occur, any diminution in the strength of that fixation as a result of feeding at the locus of shock could be ascribed either to direct interference with the fixated habit or to interference with some process maintaining it. If it were shown that feeding did not interfere directly with the habit, the hypothesis that its effectiveness in reducing fixation was the result of interference with or elimination of a maintaining mechanism (secondary reinforcement resulting from anxiety reduction) would be supported.

Experimental Procedure

Subjects

The *S*s were 96 naive male rats obtained from three colonies maintained at the State University of Iowa, as follows: 28 hooded and 28 albino rats—Psychology Department; 16 hooded rats—Anatomy Department; 24 albino rats—Pharmacology Department. The ages of the animals ranged from 57 to 75 days at the beginning of the experiment. An equal number of animals from each colony and strain was assigned at random to each experimental condition.

Apparatus

The apparatus was a single-unit T-maze constructed of unpainted white pine. Its sides were 3½ in. high; its floor plan and other dimensions were as indicated in Fig. 1. The roof of the starting box, through which animals were placed in the maze, was of colorless plexiglas, as were the roofs of the stem and the choice point. The remainder of the maze was covered with ¼ in. wire mesh. The doors were of plywood (with the exception of a plexiglas door dividing the choice point from the stem), and were operated vertically by *E* from behind a one-way screen situated in front of the apparatus. Black cloth curtains placed in front of the doors to the goal boxes prevented the *S*s from noting the presence or absence of food in the goal boxes prior to their choice.

The starting box contained an end wall which could be pushed forward to move animals into the stem; a single continuous grid covered the floor of the stem and the choice point; separate grids covered the floors of the arms from the choice point up to either goal box. Each grid consisted of two separate lengths of no. 20 copper wire wound in alternate turns ¼ in. apart about a piece of ¼ in. bakelite. The floor and walls of the left goal box were covered with fine wire screening; the right goal box was lined with cork matting. Use of these different cues was designed to accelerate learning by decreasing the generalization of reinforcement from one side of the maze to the other.

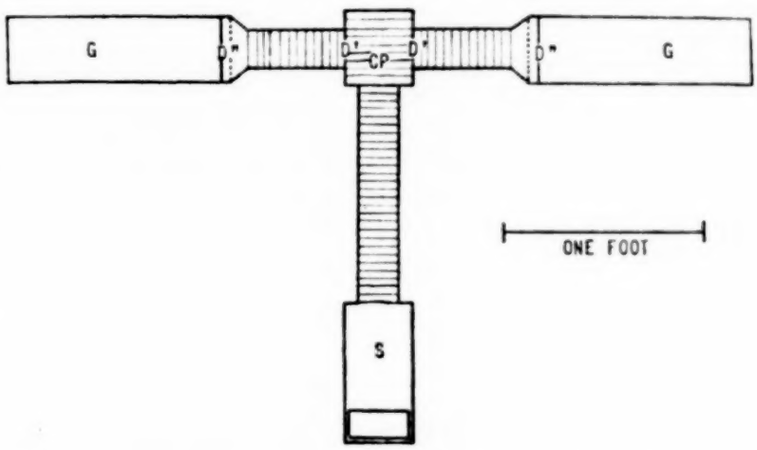

FIG. 1. Floor plan of maze. S—starting box; CP—choice point; G—goal box; D—door; dotted line in front of door—curtain; shaded area—grid

Shock was produced by means of a direct current, the source of which was a type 6C6 pentode vacuum tube provided with a plate voltage of 650 volts. The magnitude of the current was adjusted by means of a variable bias voltage from a battery and its associated potentiometer. The grids could be independently activated by means of three single-throw knife switches. Variations in the bodily resistance of animals in series with this circuit produced no measurable change in the current.

Preliminary training

Prior to training all animals were given five days of habituation to the feeding regimen of the experiment; each received eight gm. of Purina Dog Chow in individual feeding cages once daily at the hour in the afternoon at which they were to be run. Three days of preliminary training were then given as follows:

Day 1—Four trials in a straight alley (38 in. long and built of unpainted white pine) to food, consisting of a single small pellet (approximately .10 gm.) of dog chow.

Day 2—Fifteen min. of exploration in the T-maze in groups of two or three, followed by four trials in the straightaway, through a black cloth curtain, to food.

Day 3—Four non-reinforced trials in the T-maze, consisting of two runs to the right and two to the left. The first two runs were free; runs on trials 3 and 4 were forced when necessary. This procedure was followed in order to reduce to some extent any strong position preferences.

Method in learning series

Beginning on day 4 all animals were given 10 trials a day for the remainder of the experiment. At the beginning of each day's runs the animals were under a 20 to 21-hour hunger drive. The inter-trial interval was always at least 10 min. and the modal interval 15 min. On trial 1 of the learning series food was present in both goal boxes. An animal's choice on this trial was assumed to define its preference and thereafter food was presented only on that side. Neither correct nor incorrect responses involved any delay; runs to the incorrect side were unreinforced either by the presence of food or by the presence of the food-container (a metal tray two inches square). Correction, i.e., retracing after an incorrect response, was not allowed.

After 40 trials *Ss* were assigned at random to two experimental groups and to two control groups, and were then run for an additional 60 trials. The control animals continued under the same conditions as before; but the experimental animals, though continuing to receive food on their preferred side, were now administered a shock after choice [4] by means of the grids located in the arms of the maze. The grids on both sides were activated, so that these animals were shocked whether or not they continued their old response, i.e., the shock could not be avoided merely by a reversal of choice. On the first three trials shock was administered just as the animals were leaving the grid and entering the goal box; thereafter they were shocked as soon as they touched the grid. The shock was progressively increased during the shock trials from an initial intensity of 50 microamperes to a final intensity of 140 microamperes.

Since the experimental animals soon refused to approach the shock grid it was necessary to administer an occasional shock in the stem of the maze. The time allowed for a choice before shock was given in the stem was progressively reduced during the course of the shock trials from 10 to three sec. This served roughly to equate the running times of the experimental and control animals and possibly influenced the strength of habits in the experimental group.

During the shock trials the end wall of the starting box was moved forward at the beginning of each trial (for the control as well as for the experimental groups). Consequently, the shock could not be avoided by refusal to leave the starting box nor escaped by running back into it.

Maze-feeding procedure

As indicated above, both the experimental and control animals were divided into two groups. Half of the experimental animals and half of the control animals were fed in the maze immediately after the last trial of the learning series; the other half of these groups were not given this experience. The feeding procedure was as follows: the door separating stem from choice point, the door (D′) leading to the incorrect alley and the door (D″) leading to the correct goal box were closed. The area thus circumscribed consisted of the choice point (4 in. × 4½ in.) and the arm leading to the correct goal box (approximately 2 in. × 7½ in.). Two grams of food were scattered on the grid, the roof above the choice point was removed, the animal was placed directly on the grid facing in the direction of the goal, and the roof was then replaced. The animal was allowed to feed for 10 min. and was then removed. Two such feeding sessions about 1½ hours apart were given. The maze was then carefully cleaned in order to eliminate any traces of food.

At the end of this time the animals not fed in the maze were given their usual daily food ration and the others received an amount which, added to the food eaten in the maze, equalled their daily ration.

The four subgroups will be designated as follows: SF, shock animals fed in the maze before extinction; S, shock animals not fed in the maze; NSF, non-shock animals fed in the maze before extinction; NS, non-shock animals not fed in the maze. The shock animals when considered as

[4] In order to investigate the effectiveness of anxiety elimination as a means of preventing fixation, it was first necessary, of course, to determine the specific conditions under which fixation could be clearly established. The use of shock after choice as well as a number of other procedures described in this section was decided upon only after rather extensive preliminary investigation. It is by no means believed that the finally-adopted procedure is the one calculated to produce the greatest possible degree of fixation under shock conditions. This procedure, however, proved effective. The fact that certain other procedures, some of which have been favorably reported elsewhere, proved less effective emphasizes the necessity for systematic investigations designed to determine the specific variables of which strength of fixation under shock conditions is a function.

a group will be referred to as the experimental group and the non-shock animals as the control group.

Method in extinction series

On the following day, approximately 21 hours after the last trial of the learning series, extinction was begun. The food reward was shifted to the non-preferred side for each animal, and no shock was administered. All animals were run 10 trials a day under relatively distributed practice conditions (10 min. or more between trials) to a criterion of two successive responses to the non-preferred side, which now led to food.

Results

Learning series

Learning scores for the initial 40 trials of the learning series are presented in Table I. Examination of these data reveals that the differences among the four subgroups during these trials were negligible.

TABLE I

Mean Number of Correct Responses in Four Groups during the First 40 Trials (Food Reinforcement only) of the Learning Series (N = 24 in each group)

	Experimental Group		Control Group	
	S	SF	NS	NSF
M	38.50	38.71	39.08	38.25
SD	2.08	1.72	1.25	2.83

During the last 60 trials of the learning series, the experimental groups (S and SF) were shocked whereas the control groups (NS and NSF) were not. The mean number of correct responses made by each subgroup during these trials are presented in Table II. As in

TABLE II

Mean Number of Correct Responses in Four Groups during the Last 60 (Shock) Trials of the Learning Series (N = 24 in each group)

	Experimental Group		Control Group	
	S	SF	NS	NSF
M	58.29	58.54	58.71	58.38
SD	3.15	1.80	1.57	1.73

the first 40 trials the mean total scores show no significant differences among the subgroups (the between-groups variance being smaller than that within groups).

However, when comparisons were made between the experimental and control groups during shock trials on a day by day basis instead

of on the basis of total correct score, a somewhat different picture was presented. The results of such a comparison are shown in Fig. 2, which gives the percentage of errors made each day by the experimental and control *Ss* throughout the learning series.

It may be seen from these curves that during the period in which both groups were treated identically (first 4 days) their percentages of errors were highly similar from day to day. However, during the shock period (last six days) the error curves differed considerably. Introduction of shock on day 5 for the experimental group was accompanied by a sharp increase in errors for that group; [5] after 20 shock trials this curve drops to a point somewhat below that of the control group.

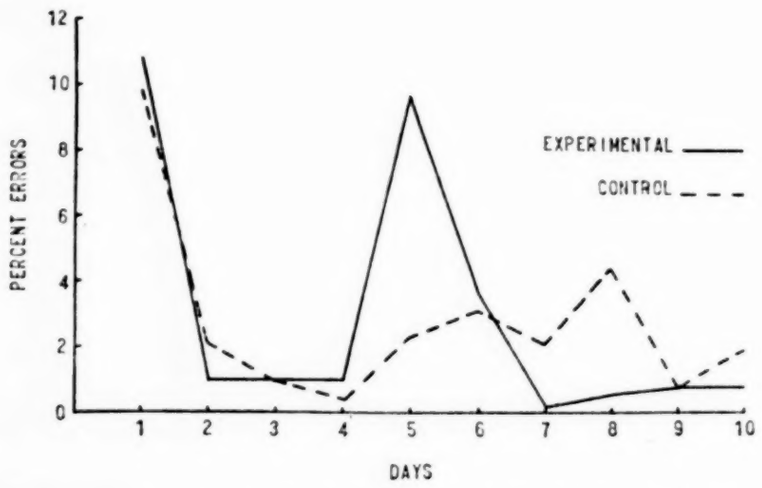

FIG. 2. Percentage of errors made on successive days of learning series by experimental and control groups ($N = 48$ for each group). Days 1–4 involved identical treatments for both groups; days 5–10 involved shock for the experimental group, no shock for the control group.

These results appear to be of some interest in relation to both regression theory and to fixation as defined in terms of response consistency. It is seen that although the error curve of the experimental group rose sharply when shock was introduced, the disruption of the on-going responses in these *Ss* was only temporary. In three animals whose results are not included in the present discussion shock effected a more permanent disruption of the on-going responses, so that these animals continued to run incorrectly throughout the shock trials.[6]

[5] An increase in errors for the control group along with some variability from day to day during the shock trials should be noted. These phenomena do not ordinarily occur under non-shock conditions in this type of learning situation. In this instance it is probably attributable to the fact the control and experimental *Ss* were run concurrently. Cleaning of the grids following the occasional urination of the shock animals was observed to disrupt to some extent the behavior of the non-shock animals subsequently run in the maze.

[6] The subsequent behavior of these animals during the extinction series is believed to have important implications for a general theory of fixation. However, since the present study is concerned only with the behavior of the animals which continued to run correctly, the data for these *Ss* have been omitted.

Since the very large majority (48) of the shock animals showed no such disruption of behavior, previous findings with respect to the strong perseverative tendencies of strong habits even when shock is introduced are strongly supported. The additional fact that the mean number of errors during the shock trials was so similar in the experimental and control groups indicates that fixation defined in terms of response persistence may be an insensitive measure of habit strength when frequency of response nears 100 percent. In the present experiment the on-going habit represented an initially preferred response which was further reinforced by food for 40 trials. It was scarcely to be expected that the two groups would exhibit wide differences in variability thereafter. The inadequacy of response consistency as a differential measure of fixation in this instance does not, of course, invalidate its use under other conditions.

Extinction series

The speed with which the original habit was extinguished under the four conditions is revealed in Table III, which gives the number of trials required by each *S* to reach the criterion of two reversals and the medians, means and SD's of the extinction scores for each group.

In order to test the hypothesis that response fixation, defined in terms of relative resistance to extinction, may result from the reduction of anxiety, it was first necessary to demonstrate that under shock there was greater fixation than under the non-shock condition. Examination of the data in Table III shows that this condition was satisfied. The S animals, it will be seen, required on the average many more trials to reach the extinction criterion than did the NS animals. The extent of the difference is indicated not only by the measures of central tendency but also by the relatively small amount of overlapping between the two distributions. Thus, the largest score (17) for the NS group was below the 30th percentile score for group S, while the lowest score for the latter was at the median of the NS group. There can be little doubt that shock, as compared with absence of shock, was effective in fixating the originally learned response.

If our major hypothesis with respect to the relationship of anxiety to fixation was correct, the elimination of the anxiety should have reduced the tendency to fixation. It was supposed that the feeding of the *S*s in the anxiety-arousing situation would have this effect. We therefore turn to a comparison of the extinction scores of the experimental animals which were subjected to this 'therapeutic' treatment (feeding in the maze) with those which were not.

Perhaps the most convincing proof that feeding at the locus of shock did reduce fixation, i.e., decrease resistance to extinction, is

TABLE III

NUMBER OF TRIALS TO EXTINCTION FOR ALL Ss IN FOUR SUB-GROUPS
(N = 24 in each group)

	S	SF	NS	NSF
Psychology Hooded	11	17	8	10
	14	15	17	6
	12	3	7	12
	8	7	17	8
	57	7	5	10
	41	10	8	7
	48	4	7	5
Psychology Albino	19	10	6	7
	25	5	9	13
	63	26	6	19
	14	16	8	6
	24	21	10	8
	27	4	11	9
	33	4	8	4
Pharmacology Albino	255*	19	5	7
	263*	3	15	2
	15	6	6	3
	57	33	17	5
	33	3	17	9
	163	98	11	6
Anatomy Albino	36	10	11	17
	193	3	8	28
	10	10	7	17
	46	27	11	10
Mdn.	33	10	8	8
M	61.12	15.04	9.71	9.50
SD	74.09	19.24	3.83	5.75

* These two animals refused to extinguish at all under the ordinary extinction procedure. After 25 days during which they ran 250 trials without a single reversal of their original habit their inter-trial interval was reduced to two min. Under this relatively massed condition, they reached the criterion within a short time.

furnished by examination of the individual scores. Thus, it is seen that only two of the 24 members of the SF group had as large a score as the median animal in the S group; and conversely, only two members of the S group had as small a score as the median animal in the SF group. The differences between the measures of central tendency were strikingly large—almost as large, in fact, as those between the S and NS groups.

A *t*-test of the significance of the difference between the means of groups S and SF was not possible, since their variances differed greatly and the distributions departed markedly from normality. However, a number of other types of tests of significance were applicable.

Chi-square was employed in testing the hypothesis that the two groups were homogeneous with respect to the number of responses required for extinction. Classifying the number of animals in each group into those which required less than 15 trials to extinguish and those which required 15 or more trials, the following contingency table was obtained: [7]

No. Trials to Extinction	S	SF
Below 15	6	15
15 or above	18	9

The χ^2 value for these data is found to be 6.86, which, for one degree of freedom, is significant beyond the one percent level of confidence.

If the results are analyzed by means of a technique suggested by Festinger (**2**) for determining the significance between means of samples drawn from positively skewed populations, it is found that the ratio between the means yields an *F*-value of 4.53, which for the degrees of freedom involved (33 and 29), is significant well beyond the one percent level.

A further test of the mean difference between these groups was obtained by first transforming the original extinction scores into log log scores. This transformation satisfied the requirements of normality and equality of variances in the two sets of scores and thus permitted the use of the *t*-test. The resulting *t* is 4.87, which, for 46 d.f., is well above the value required for significance at the one percent level of confidence.

The agreement among the results of these tests of significance afforded a very sound basis for rejecting the null hypothesis (that the groups did not differ in their extinction scores) and presented a convincing verification of our prediction that feeding at the locus of shock would reduce the tendency to fixation. There was reason to believe, in fact, that the feeding procedure not only reduced the degree of fixation in the SF group but *on the average* prevented fixation entirely. When the extreme score (98) in the SF group was omitted so that the conditions for the application of the *t*-test could be met, no basis was found for rejecting the hypothesis that the SF group did not differ from the two non-shock groups. Failure to obtain such a difference is believed to lend even stronger support to the interpretation that fixation is, in the present experiment at least, closely related to secondary drive and its associated reinforcing mechanism.

[7] The value 15 represents the mean number of trials to criterion for SF animals. Other bases for dividing the groups might have been used. The distributions overlap so little, however, that it is difficult to find a score for this purpose which has some ordinary statistical significance and also permits conformity to the stricture that the theoretical frequency in each cell be greater than ten. For instance, use of the median of either group as the point of division results in higher values of χ^2, but does not allow a test involving theoretical frequencies greater than 10.

It has been proposed that this prevention of response fixation in the SF group was the result of interference with the mechanism of anxiety reduction which maintained the response, rather than the result of direct interference with the original response. This latter alternative was tested by comparing the resistance to extinction of the NS and NSF animals. If the procedure of feeding in the maze directly affected the original response, weakening it in some manner, the means of these two groups should have differed. A *t*-test of the difference between the means yields a value of .15, which for 46 d.f. would be obtained by chance 80 to 90 percent of the time. The results therefore failed to support this alternative hypothesis and thus added credibility to our supposition that the decreased resistance to extinction of the SF group is to be attributed primarily to the elimination of anxiety through feeding.

Influence of colony differences on fixation

Among the important variables found to be related to the strength of fixation of shock animals was the particular colony from which they were obtained Examination of the individual scores in Table III reveals that there were large differences between the animals bred in the Psychology laboratory and those obtained from the Anatomy and Pharmacology laboratories.

TABLE IV

MEANS AND SD'S OF PSYCHOLOGY AND NON-PSYCHOLOGY ANIMALS FOR FOUR EXPERIMENTAL CONDITIONS

	Psychology Animals (N = 14 for each condition)		Pharmacology and Anatomy Animals (N = 10 for each condition)	
	M	SD	M	SD
S	28.29	17.11	117.10	92.82
SF	10.64	6.98	21.20	27.48
NS	9.07	3.58	10.80	4.17
NSF	8.86	3.73	10.40	7.64

Table IV presents the means and SD's of the distributions of scores in the four sub-groups for the Psychology and non-Psychology animals separately. The differences between the hooded and albino strains in the Psychology group were negligible, and the relatively small N's for each of the other groups did not warrant their separation for purposes of statistical analysis. Consequently, all Psychology animals were combined into one group, while all others were combined into the second.

It is apparent that the most strongly fixated animals were those from Pharmacology and Anatomy. It might then be questioned

whether the difference obtained between the total S and SF groups was actually restricted to the non-Psychology animals and not found in the Psychology *Ss*. The *t*-test applied on the hypothesis that there was no difference between the S and SF animals from Psychology yields a value of 3.44, which, at 26 d.f., is well beyond the one percent level. Thus it is seen that a significant difference in the fixation of the S and SF groups was obtained in this more limited group of *Ss*.[8]

The fact that the average strength of fixation in the non-Psychology *Ss* far exceeded that of the Psychology animals may be of no little significance for our theory. It will be recalled that our major hypothesis states that when fixation occurs as a result of shock, it is due to the reinforcing effect of anxiety reduction. Presumably, therefore, the degree of fixation is a function of the strength of the anxiety drive. Since anxiety is thought to be essentially a visceral response, the general emotionality of individual animals might be taken as an index of the amount of anxiety they would show in a given situation and consequently of the extent to which they would be likely to maintain non-adaptive responses in that situation.

Note was made on alternate days of the learning series of all animals that defecated during the experimental period. Inasmuch as defecation is commonly regarded as a sign of emotionality, the number of Psychology and non-Psychology *Ss* respectively that defecated during the experiment (particularly during the initial learning series, which involved only a routine learning situation) could be taken as a rough measure of the anxiety level of the two groups.

The data showed that defecation occurred in 20 of the 56 Psychology animals and in 33 of the 40 non-Psychology animals. The χ^2 test of homogeneity based on these frequencies yields a value of 20.6, which, for one d.f., is significant well beyond the one percent level. If defecation is accepted as a valid measure of emotionality, one may conclude that the non-Psychology *Ss* were more emotional than those from Psychology.

There is considerable reason to believe that the Psychology animals represented a population which has resulted from selective breeding for non-emotionality. Animals in psychological research that are difficult to handle, which refuse to run and which cannot learn maze and discrimination problems are usually discarded. Those used as breeding stock, therefore, are precisely the ones which make a more rapid adjustment to new situations, i.e., are less emotional. This is presumably not the case with animals bred in laboratories in which behavior problems are not studied. In consequence, it was not surprising that Psychology animals showed less anxiety and there-

[8] The difference of 95.90 between the means for the S and SF groups in the other animals was found, as might be expected, to be highly significant by the Festinger test.

fore fewer fixation tendencies than other animals, under the conditions of the present experiment.

Discussion

One of the bases for the belief that fixation is not to be explained in terms of learning principles appears to be the fact that *any* response elicited by shock tends to be fixated, *regardless of its strength prior to the time shock is introduced.* It is not difficult, certainly, from the point of view of learning theory, to explain why any response, *once it has occurred*, might persist during shock trials. Despite the fact that it may never have been previously reinforced, or may not currently represent an efficient habit as far as immediate access to food is concerned, the fact that it leads to escape from shock would nevertheless be an adequate basis for predicting its continuance. We should, however, examine more carefully the grounds for the claim that the likelihood of evocation of a particular response by shock is independent of its relative strength during the pre-shock period.

Among the results tending to support this conclusion are data reported by Everall (1). She found that only one animal in ten shifted to a new response during shock trials after having learned a previous response to a criterion of 38 correct out of 40 trials, whereas two animals in ten shifted after having learned a previous response to a criterion of 111 out of 120 trials. In this instance it appears likely that the strength of responses learned to the lesser criterion generally sufficed to preclude successful competition of another response. It was not to be expected, therefore, that further training on these responses should make an appreciable difference.

It is of interest in this connection to compare the results obtained by Everall with her 40-trial food-reinforcement group with those of the present experiment, which also involved 40 food-reinforced trials in the pre-shock period. Despite certain procedural differences which tended to minimize response shifting in our experiment, the percentages of animals showing disruption of response with shock were very similar, being 10 and six percent respectively, in the two experiments. It may be concluded that in these situations the tendency to disruption of response (shifting to a response not previously reinforced) was slight indeed. Everall herself decided on the basis of additional studies that "when shock was administered before a choice point, the majority of animals reacted in the direction they had selected most frequently in the past" (1, p. 357).

The results of a number of other experiments indicate that the relative strengths of habits have a decided influence on the tendency for those responses to be elicited in a shock situation. Thus, Mowrer (14) and O'Kelly (19) found that animals continued their correct

response when shock was introduced *if no other response had been previously learned*, whereas animals which had previously learned a different response tended to revert to it. Martin found that the greater the relative amount of training on an earlier habit, the less the tendency to persevere in a later habit during shock; he concluded that "other things being equal the amount of training on the responses involved will be the decisive factor in the occurrence of regression, and that native or unlearned preferences have no effects that cannot be counter-balanced by training" (**10**, p. 13). Kleemeier (**6**), who does not favor the view that the relative strengths of habits are an important factor in determining response to shock, nevertheless found that a much larger proportion of animals responded to shock by running to an alley which had been previously learned than by running to an alley which had not been previously learned. The weight of the evidence thus offers little support to the interpretation that the factors operating to produce responses under shock conditions are independent of those which determine responses in other ordinary learning situations.

A second basis for the belief that fixation may represent an abnormal phenomenon has been the observation that fixated animals may show considerable constriction of behavior. Thus, in Kleemeier's investigation it was found that when their fixated response was blocked, a great many animals refused to run altogether. This behavior, according to Kleemeier, is not to be explained on the assumption that shock choices are determined by their relative habit strengths. He writes: "If this were the determining factor then elimination of [the alley with] the highest excitatory value should merely cause a shift in behavior toward the alley with the next highest value" (**6**, p. 32).

If the word 'response' had been substituted for the word 'alley' in this quotation, the position of the proponents of a learning interpretation would have been fairly stated. If one examines the conditions under which ten out of 16 animals in Kleemeier's first experiment finally refused to run, one finds that this response was actually fairly low in the response hierarchy. It was elicited in these animals only after a long sequence (varying somewhat for two sub-groups) which included, during some 200 trials, the following order of events: (1) training in two different habits; (2) shock at the choice point; (3) blocking of any new responses elicited by the shock; (4) introduction of delay of reward for the resulting habit; (5) removal of shock; (6) introduction of shock after choice; and finally (7) blocking of the then on-going habit. It is noteworthy that during this somewhat elaborate series of conditions, seven of the 16 animals experienced blocking of the alley to which they had initially responded when

shocked and all animals experienced blocking of the alley which they eventually fixated. It is scarcely remarkable, under these circumstances, that 10 of the 16 *Ss* should have refused to enter any other alley. From the point of view of learning theory, one would have to consider the effects of such blocking, particularly the likelihood of their generalization to the other alleys in the maze. Such generalized extinction effects might be expected [9] to produce just such failure of response as Kleemeier found.

A final basis for the belief that fixation represents a non-learning phenomenon is its apparently unusual (abnormal?) strength. Thus animals may persevere in a response elicited under shock conditions for what may appear to be an 'unreasonable' length of time. For instance, in Kleemeier's study, a comparison of the effect upon resistance to extinction of 50 food reinforced trials with that of a like number of shock trials revealed a truly impressive difference. Whereas no animal in his non-shock group required more than 12 trials in which to discard a non-adaptive response (involving delay of reward), 15 percent of his shock *Ss* failed to discard the response in 50 trials.

Comparable results were obtained in the present study. When the on-going response was made incorrect (no longer leading to food) none of the non-shock animals required more than 28 trials to reverse this response, despite the fact that it had been initially preferred and had been further reinforced for 100 trials by food reward. In the shock (S) group on the other hand, 29 percent of the animals persisted in the response for as many as 50 trials and 17 percent for as many as 150 trials, during which they went entirely unrewarded as far as food was concerned.

It is this phenomenon of persistence of response in the absence of any apparent reward (goal) which seems most contrary to learning theory. Learned behavior, according to the widely accepted view, is goal directed, whereas fixated behavior appears to be merely repetitious. It seems to continue without rhyme or reason, despite the fact that it may no longer be effective in escaping shock or reaching food or, in fact, achieving any apparent goal.

We have proposed to explain such fixation in terms of principles derived from ordinary learning experiments. Assuming that the pain response (anxiety) may be conditioned to the cues contiguous with the noxious stimulus and that escape from this anxiety constitutes a reinforcing state of affairs, it is possible to explain many of the findings with respect to response fixation under shock conditions.

[9] Cf. Hull (4), who deduces that "in certain situations the organism will give up seeking, i.e., cease making attempts, and thus fail to perform the correct reaction even when it possesses in its repertoire a perfectly correct excitatory tendency" (p. 13).

It is important to note that the present treatment *supplements* an explanation of fixation in terms of habit strength alone. Although it is believed that escape from shock ordinarily constitutes a very strong reinforcement and probable results in habits of great strength, it would be possible under the present set of postulates to conceive of the fixation of a relatively weak habit so long as a *maintaining mechanism* resulting from anxiety reduction were operating. This view appears to be most consistent with the common clinical observation that mal-adaptive habits may be readily extinguished *if the anxiety aroused by the situations in which they occur is eliminated first.*

In view of the successful prediction of the results of the present study on the basis of the assumption of secondary reinforcement (and the likelihood that the same mechanism has operated as an uncontrolled variable in a number of studies relevant to this problem), it would seem appropriate to conclude that the continuation of acts under conditions which might be expected to result in their extinction is, in fact, due to factors operating in ordinary learning situations. This possibility must be given serious consideration before the view is accepted that fixation is not explicable in terms of learning concepts and principles.

Summary

The present study was concerned with the problem of the extent to which the persistence of a non-adaptive response in a shock situation may be the result of uncontrolled secondary reinforcement resulting from anxiety reduction. The anxiety reduction hypothesis involves the following assumptions:

1. Intense stimuli such as shock evoke in the organism certain responses, mediated largely by the sympathetic division of the autonomic nervous system, which produce an internal drive state.
2. External cues present at the time of shock acquire the capacity to evoke these visceral responses and thereby the drive, which, when elicited in this manner, is termed 'anxiety.'
3. Removal of these cues reduces the anxiety and therefore constitutes a reinforcing state of affairs.

It follows from these assumptions that even after shock has been removed, the secondary reinforcement resulting from the reduction of anxiety following escape from the cues eliciting it may materially retard the extinction of any response closely associated with these cues.

Among the implications of this theory is the consideration that the elimination of the anxiety in this situation should preclude the possibility of maintenance by means of secondary reinforcement alone of responses which are no longer adaptive. It was hypothesized that

feeding in a situation in which anxiety had been developed would be adequate to remove the anxiety state, thus eliminating this source of secondary reinforcement and decreasing the resistance to extinction of any non-adaptive response.

To test these implications, two groups of 24 rats were given 100 trials in a single-unit T-maze, with food reward in the goal box on their preferred side. During the last 60 trials these animals were shocked immediately after the choice point in the maze. At the completion of this training, one group was fed in the maze at the locus of shock for two ten-minute periods. On the day following this 100-trial learning series, the food reward was shifted to the goal box on the non-preferred side for each animal, and no shock was administered. All animals were run until their original response was extinguished, and the number of trials to extinction for the shock (S) animals and for the shock-fed (SF) animals were compared.

As a control for the possibility that the feeding activity affected the original responses directly rather than by way of anxiety elimination, two additional groups of 24 *S*s each were trained under the conditions indicated above, but were not shocked. As in the case of the shock animals, one group (NSF) was fed in the arm of the maze immediately after the choice point, while the other group (NS) was not. These groups were then compared on the basis of the number of trials required to extinguish the original response.

The following results were obtained:

1. Responses of the S animals showed great resistance to extinction as compared with those of the NS animals.
2. Responses of the S animals were significantly more resistant to extinction than those of the SF animals.
3. There was no significant difference in resistance to extinction between the responses of the SF animals and those of the two control groups.
4. There was no significant difference in resistance to extinction between the responses of the two control groups.

It was concluded that:

1. The experimental conditions were adequate to effect fixation in the animals run under shock conditions.
2. The feeding procedure was 'therapeutic' in that it prevented fixation in the average SF animal.
3. The prevention of fixation in the SF group was the result of interference with some mechanism maintaining the fixated responses rather than the result of direct interference with the responses themselves.

On the basis of these considerations, the inference has been drawn that fixation resulting from shock may be the result of the operation of secondary reinforcement resulting from anxiety reduction, and therefore due to factors operating in ordinary learning situations.

(Manuscript received April 7, 1947)

References

1. Everall, E. E. Perseveration in the rat. *J. comp. Psychol.*, 1935, **19**, 343–369.
2. Festinger, L. A statistical test for means of samples from skew populations. *Psychometrika*, 1945, **8**, 205–210.
3. Hamilton, J. A., & Krechevsky, I. Studies in the effect of shock upon behavior plasticity in the rat. *J. comp. Psychol.*, 1933, **16**, 237–253.
4. Hull, C. L. Mind, mechanism, and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1–32.
5. Klee, J. B. The relation of frustration and motivation to the production of abnormal fixations in the rat. *Psychol. Monogr.*, 1944, **56**, 1–45.
6. Kleemeier, R. W. Fixation and regression in the rat. *Psychol. Monogr.*, 1942, **54**, 1–34.
7. Maier, N. R. F. The specific processes constituting the learning function. *Psychol. Rev.*, 1939, **46**, 241–252.
8. Maier, N. R. F., Glaser, N. M., & Klee, J. B. Studies of abnormal behavior in the rat: III. The development of behavior fixations through frustration. *J. exp. Psychol.*, 1940, **26**, 521–546.
9. Maier, N. R. F., & Klee, J. B. Studies of abnormal behavior in the rat: XII. The pattern of punishment and its relation to abnormal fixations. *J. exp. Psychol.*, 1943, **32**, 377–398.
10. Martin, R. F. 'Native' traits and regression in rats. *J. comp. Psychol.*, 1940, **30**, 1–16.
11. McClelland, D. C. Functional autonomy of motives as an extinction phenomenon. *Psychol. Rev.*, 1942, **49**, 272–283.
12. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale University Press, 1941.
13. Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, **46**, 553–565.
14. Mowrer, O. H. An experimental analogue of 'regression' with incidental observations on 'reaction formation.' *J. abnorm. soc. Psychol.*, 1940, **35**, 56–87.
15. Mowrer, O. H. Anxiety-reduction and learning. *J. exp. Psychol.*, 1940, **27**, 497–516.
16. Mowrer, O. H. The Freudian theories of anxiety: a reconciliation. 1941. Address given before Monday Night Group of the Institute of Human Relations.
17. Mowrer, O. H., & Jones, H. Habit strength as a function of the pattern of reinforcement. *J. exp. Psychol.*, 1945, **35**, 293–311.
18. O'Kelly, L. I. An experimental study of regression: I. The behavioral characteristics of the regressive response. *J. comp. Psychol.*, 1940, **30**, 41–53.
19. O'Kelly, L. I. An experimental study of regression: II. Some motivational determinants of regression and perseveration. *J. comp. Psychol.*, 1940, **30**, 55–95.
20. Sanders, M. J. An experimental demonstration of regression in the rat. *J. exp. Psychol.*, 1937, **21**, 493–510.
21. Sears, R. R. Survey of objective studies of psychoanalytic concepts. *Bull. Soc. Sci. Res. Coun.*, 1943. No. 51.
22. Tolman, E. C. *Drives toward war.* New York: D. Appleton-Century, 1942.

AN INVESTIGATION INTO THE CAUSES OF RETROACTIVE INTERFERENCE

BY CHARLES E. OSGOOD

University of Connecticut *

Introduction

Transfer and interference in the successive learning of verbal materials by human *Ss* have been studied extensively, many recent investigations in the field being directed toward discovering the causes of retroactive interference.[1] The work of McGeoch and his associates (**3**, **6**, **7**, **8**) has offered evidence for a *competition of response theory* which originates in experimental evidence showing retroactive interference to increase with the degree of similarity between original and interpolated materials. According to this view, a response is interfered with when another takes its place, overtly or implicitly. Gibson's *generalization hypothesis* (**1**) refines this theory and formulates it in conditioning terms, rendering more specific and predictable the conditions under which more or less interference may be expected. Melton and others (**9**, **10**, **20**) have demonstrated that part of the total interference must be attributed to factors other than competition at the point of recall, one such factor being tentatively identified as *unlearning* of the original responses during interpolation.

Competition and unlearning processes are combined in what is now known as the *two-factor theory*, in which one factor is the competition of original and interpolated responses at the point of recall and the other is the unlearning of the original responses during interpolation. This is still a competition theory of retroaction, since the effect of unlearning is to reduce the effectiveness with which the original responses compete at the time of recall. One hypothesis arising from the two-factor theory is that proactive interference (PI) should be less than retroactive interference (RI), and this has been verified in two investigations (**8**, **10**). Another hypothesis, arising from the observation that unlearning is restricted largely to the first few inter-

* This research was conducted in the Psychological Laboratory of Yale University, and the writer wishes to express his gratitude for the time and facilities provided there. The writer also wishes to express deep gratitude to Davis H. Howes who served not only as experimenter but also as intellectual catalyst.

[1] Interference is used here rather than the more usual term, inhibition. The term actually refers to an observed decrement in learning or recall and not to a process. The observed decrement may or may not be due to inhibition. Henceforth in this paper the term *interference* will mean observed decrement and the term *inhibition* will be reserved for the psychological process.

polated trials, is that the difference between PI and RI should be constant beyond the first few trials, and this has also been verified (**20**).

In terms of this evidence, the two-factor theory appears adequate to incorporate the relevant phenomena. However, a recently published article by the writer (**12**), using meaningful materials in the traditional retroactive interference paradigm, found similar interpolation to result in *less* interference than neutral or opposed materials. Since the similarity of responses is presumed to determine the degree of competition among them and hence the amount of RI, this finding casts doubt upon the adequacy of the two-factor theory. Furthermore, in this study it was found that, although similar materials produced less total interference, these materials showed the greatest frequencies of interlist intrusions, for both proactive and retroactive situations; thus neither unlearning nor competition factors, as indexed by such intrusions, were capable of explaining the difference in total RI.

The present study duplicates the general design used by Melton and Irwin (**9**) in that, after equated learning of the original materials, varying degrees of interpolated learning are given to different groups of subjects. But, whereas Melton and Irwin used neutral nonsense materials, the present investigation compares the interference effects of similar and opposed meaningful materials. Significant differences are found between these two types of material, both in terms of total proactive and retroactive interference and in terms of the relative frequencies of blanks and intrusions, and the two-factor theory does not incorporate the findings.

Method

Since essentially similar general procedures and apparatus were employed in a previously reported study (**12**), the reader is referred to that source for further details. In the present experiment sets of 14 paired-associates were learned by 80 *S*s, students in the writer's introductory psychology courses, by the anticipation method in which the stimuli were letter-pairs such as *t.g.* and the responses were common adjectives. The traditional retroactive interference paradigm was followed (A—B; A—K; A—B), in which one experimental variable was the meaningful relationship between B and K adjectives (similar or opposed in meaning) and the other variable was the degree of interpolated learning. The experiment was so designed that each *S* learned interpolated materials having seven words in each meaningful relation to the original materials, and each word in the original learning had interpolated (with an equal number of *S*s) a meaningfully similar and meaningfully opposed word. This balance of materials was preserved identically for all degrees of interpolated learning; one group of *S*s practiced to a criterion of one correct anticipation of 10 out of the total 14 words, another group of *S*s to a criterion of two correct anticipations, a third group to a criterion of four correct anticipations, and a final group practiced to a criterion of eight correct anticipations of 10 of the total 14 words. There were 20 *S*s in each group. In this manner individual *S*s and individual words serve as their own controls for comparisons between similar and opposed conditions. The other variable, degree of interpolated learning, is controlled for the selection of learning materials but not for the selection of *S*s.

I. *Materials.*—The adjectives used in the experiment were within the functioning vocabulary of the average college sophomore and were of roughly equal length and familiarity. The words selected by *E* as similar and opposed in meaning to the original list words were the same as used

in the earlier experiment, and *E*'s selection was validated with high consistency by a group of 20 judges at that time. Words called *similar* in meaning to the original words are not synonyms (by the criterion of being listed as such in dictionaries of synonyms and antonyms), but they do have the same feeling-tone and are to a large extent interchangeable descriptively, as for example *dejected-low*, *free-open*, *weighty-serious*, and *tense-hard*. Words selected as opposed in meaning to the original words are not antonyms (by the same criterion), but they do have opposed feeling-tone, as for example *dejected-high*, *free-closed*, *weighty-humorous*, and *tense-soft*.

Table I presents all the adjectives used in the experiment in terms of their meaningful relations. Forty *S*s (10 in each interpolation group) learned one set of 14 words as original materials (O.L.$_1$), while the other half of the *S*s learned as original materials the direct opposites of these

TABLE I

ADJECTIVES EMPLOYED IN THE EXPERIMENT IN TERMS OF THE MEANINGFUL RELATIONS BETWEEN ORIGINAL AND INTERPOLATED MATERIALS

Stim.	O.L.$_1$	Interpolated Learning (I.L.) Meaningful Relations with O.L.$_1$		O.L.$_2$
		S	O	
		Meaningful Relations with O.L.$_2$		
		O	S	
c.m.	elated	high	low	dejected
f.s.	tense	hard	soft	relaxed
g.l.	free	open	closed	restricted
h.v.	dainty	clean	dirty	coarse
j.y.	robust	solid	flimsy	puny
k.t.	neat	clear	hazy	messy
d.m.	skillful	quick	slow	awkward
p.j.	distinct	sharp	blunt	blurred
q.r.	slender	airy	solid	stout
t.g.	drowsy	dead	alive	alert
v.k.	weighty	serious	humerous	spritely
w.p.	boorish	rough	smooth	gracious
x.n.	rounded	graceful	clumsy	angular
y.b.	pale	sickly	healthy	brilliant

words (O.L.$_2$). Since identical sets of interpolated materials were learned in both cases, a further control over the selection of words is provided: Each interpolated word having a similar meaningful relation to a given word on O.L.$_1$ necessarily has an opposed meaningful relation to its opposite on O.L.$_2$; thus identical words are made to function in both similar and opposed meaningful relations. For example, the interpolated word *high* is similar in meaning to *elated* (the word having the same stimulus, *c.m.*, on O.L.$_1$) while it is opposed in meaning to *dejected* (the word having the same stimulus on O.L.$_2$).

II. *Design.*—In Table I the materials are arranged in terms of their meaningful relations. The actual experiment was designed so that the interpolated materials learned by each *S* would include an equal number of similar and opposed words. This minimizes the likelihood of the *S*'s developing a 'set' for a given type of meaningful relation. The rotational design employed may best be seen by referring to Table II which shows the distribution of *S*s among the experimental conditions. Groups of five *S*s received identical treatment, learning the same original and interpolated sets as well as having the same degree of interpolated learning. Of all 20 *S*s having the same degree of interpolation (i.e., 1 IR), 10 learned O.L.$_1$ and 10 learned O.L.$_2$. Of the 10 *S*s learning O.L.$_1$, half had I.L.$_1$ and half had I.L.$_2$, the latter words being the direct opposites of those on I.L.$_1$. The 10 *S*s learning O.L.$_2$ are halved for the interpolation in the same manner.

Concrete illustrations will help clarify the nature of this design. Subject *R. S.* first learns O.L.$_2$; he then learns I.L.$_1$ to a criterion of 4 correct anticipations of 10 of the total 14 items; he then is given five relearning trials on O.L.$_2$. Five other *S*s follow this identical procedure. Following a single *item* (one stimulus and its varied response adjectives) through the design, it may be seen that whereas half the total 80 *S*s have *c.m.*—*elated* as an original paired-associate, the

TABLE II

DESIGN OF EXPERIMENT: DISTRIBUTION OF SUBJECTS AMONG LISTS OF MATERIALS AND DEGREES OF INTERPOLATION

	1 IR		2 IR		4 IR		8 IR		Total *Ss*
	I.L.$_1$	I.L.$_2$	I.L.$_1$	I.L.$_2$	I.L.$_1$	I.L.$_2$	I.L.$_1$	I.L.$_2$	
O.L.$_1$ c.m.—elated	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	40
O.L.$_2$ c.m.—dejected	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	5 c.m.—high	5 c.m.—low	40
Total *Ss*	20		20		20		20		80

other half learn *c.m.—dejected*. Of those who had *c.m.—dejected* as an original pair, half learn *c.m.—high* and half *c.m.—low* as one of their interpolated paired-associates, and the same division is true for those who originally learned *c.m.—elated*.

Thus (1) each *S* learned an equal number of similar and opposed relations, thereby serving as his own control; (2) each word on O.L. had interpolated, with an equal number of *Ss* for each degree of interpolation, a word of similar meaning and a word of opposed meaning, thereby serving as its own control; and (3) identical materials were used for all degrees of interpolated learning. The only major variable not controlled by this design is the assignment of *Ss* to the various degrees of interpolation. Since these assignments were made more or less at random, it was expected that this variable would be of no significance; in any case, it could not affect the differences between similar and opposed materials.

III. *Apparatus.*—The apparatus used in this series of studies has numerous advantages over the usual memory drum and has been described in detail in the earlier report (12). The *S* sits facing a dull-black plywood screen in which is a window wherein the learning materials are presented. The stimulus letters and response adjectives are printed in large block type on thick cards, one card for each paired-associate, and these cards are moved out along a continuous screw at a constant rate, one card dropping from the end of the screw every eight seconds. The subject can only see the front card. The presentation window is divided into left and right halves by a thin bar, the stimuli appearing on the left and the responses on the right. In the presentation of a given paired-associate, the stimulus member appears first for four sec., the right-hand half of the window being covered by a shutter; then, for a two-sec. period, the response member is revealed as well; there is then a final two-sec. period with the entire window covered by the shutter, during which time that card drops and the next in order is revealed.

This apparatus makes possible *continuous randomization of the items* during learning; this is accomplished by shuffling the item cards after each trial, thus avoiding serial effects. It is also possible with this apparatus to *equalize the degree of learning of the various items* within the set being learned; this is accomplished by dropping items from the set as they are mastered to the requisite criterion. The items still to be mastered drop from the end of the rotating screw onto a tray in the middle level of the apparatus box. When an item is anticipated correctly at the criterion level, *E* presses a foot pedal which draws this tray back and allows that card to drop through to the lowest level of the apparatus box, thus segregating mastered and unmastered items. In both the original and interpolated learning, after *S* has mastered 10 of the 14 items, *one review trial* is given with all items included, this being done to equalize as much as possible the degree of learning of the items mastered early with those more recently mastered and those still not quite learned.

IV. *Procedure.*—Following the retroactive interference paradigm, the general procedure was as follows: Each *S* first learned one set (O.L.) of 14 items to a criterion of one correct repetition; then a second set (I.L.) of 14 paired-associates was learned to a criterion of one, two, four, or eight correct anticipations; finally the *S* was given five trials on the original materials, the first

being a pure recall trial and the others necessarily involving relearning. The paired-associates were learned by the anticipation method in which the *S*, presented with the stimulus, must anticipate the correct word by making the correct response out loud before the word actually appears.

The instructions read to the *S*s in this study were practically identical with those used for the previous investigation (12) and are not quoted here. They emphasized the need for speed in responding and that guessing would not be penalized. Preceding both the original learning and the interpolation the *S*s were given one familiarization trial (on which results were not recorded) and instructed to read both the stimuli and responses out loud as they appeared, this being done only on the familiarization trial.

V. *Significant aspects of the method.*—Certain significant details in the method need clarification. (1) *The timing of the stimulus and response members* in such a way that the stimulus is exposed for four seconds was designed to permit a rough measure of the latencies (reaction-times) as a different, and perhaps more sensitive, measure of response strength than mere presence or absence of the correct response. The activation of an electromagnetic coil within the apparatus produced a sharp and distinctly audible click every two secs. If the correct response was made between the first click (appearance of the stimulus) and the second, it was scored '1'; if the correct response was made between the second click and the third (appearance of the response word), it was scored '2'; if it occurred after the third click or failed to occur entirely, a 'blank' was recorded. All incorrect reactions (intrusions) were also recorded along with their latency scores.

(2) *The degree of interpolated learning was based on the number of correct anticipations rather than the number of trials.* This method was felt to be superior in that all items are brought to roughly equal levels of learning, and this is important where the frequencies of blanks and intrusions are to be used as indexes of the causes of interference. A few poorly learned items in an eight-*trial* list may result in intrusions which obscure important trends. A correct anticipation is here termed a reinforcement, and hence the degrees of interpolated learning are symbolized as 1 IR, 2 IR, 4 IR, and 8 IR.

(3) *The selection of 1, 2, 4 and 8 interpolated reinforcements* as the degrees to be used was based on the findings of the previous study. That experiment had shown that, with a criterion of one correct repetition on the original learning and two correct on the interpolation, *S*s could make about 75 percent correct responses on the review trial of the interpolation. On this basis it was decided that four times as much interpolation (8 IR) would be adequate in the present experiment, which has proven to be the case.

(4) Although the total time required for interpolation varied considerably with the degree of interpolated learning, *no rest intervals were provided to equate the time elapsing between original learning and recall.* Rather, as soon as an *S* was through with the interpolated learning, regardless of the time required, he went directly into the relearning of O.L. This was done for two reasons. First, modern theories generally attribute forgetting to what happens in time rather than to any function of time itself, and it seems wiser, therefore, not to permit variable 'rest' periods wherein *S*'s activities are beyond experimental control. Second, Twining (19) has compared conditions under which the period between O.L. and its recall is held constant (by rest periods) with conditions where time is allowed to increase with the amount of interpolated learning; no appreciable differences were found, thus verifying the theoretical expectation.

Results

Two different measures have been utilized to indicate the degrees of proactive and retroactive interference yielded in the various conditions of this experiment. (1) *Number of correct responses*, taking into account only the presence or absence of correct responses within the time limits established, and (2) *speed of responding*, using latency (reaction-time) as an index of habit strength. In the former case interference is indicated by a decrease in the number of correct responses; in the latter case interference is indicated by an increase in the latency (i.e., slowing down) of responses. The latency measure was used in the previous study in this series (12) with satisfactory

results. Recently Postman (**13**) has evaluated reaction-time as a measure of retroactive interference, finding it to be a sensitive measure which did not, however, correlate directly with number of correct responses. In the present experiment, speed of responding was determined in the following manner: For a given trial under a given condition (similar or opposed materials) and for a given degree of interpolation, the *percentage* that '1' correct responses (i.e., fast reactions) are of all correct responses (i.e., fast and slow reactions combined) was computed, this percentage indexing the mean speed of responding. The larger the percentage, the greater the average speed of respond-

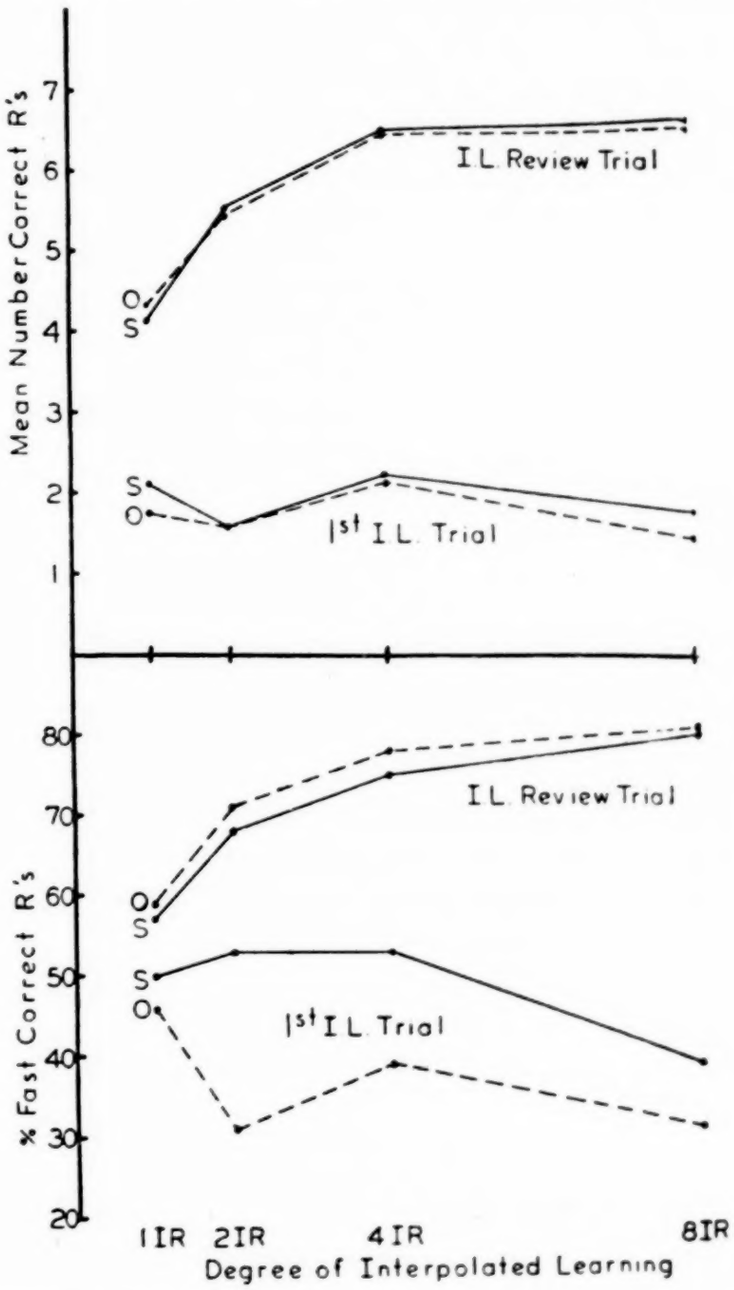

FIG. 1. Interpolated learning: mean number of correct responses (upper) and percent fast correct responses (lower) for 1st I.L. trial and I.L. review trial

ing. It should be noted that this measure takes into account *only* the latency factor, being based solely upon correct responses.

In the following section on results, *the major findings will be put in italics*, such portions serving as a summary of the section.

I. *Interpolated learning.*—Mean number of correct responses (upper portion) and mean speed of responding (lower) are given in Fig. 1 for both the 1st I.L. trial and the I.L. review trial, results for the similar condition being indicated in smooth lines and those for the opposed condition in dashed lines. The significance of differences between similar and opposed materials is estimated by the *t*-values shown in Table III. In computing the significance of these differences, both here and in the retroactive situation, it was possible to employ a formula which takes into account the inter-subject variance.[2] It will be recalled that in the design employed each *S* learned an equal

TABLE III

INTERPOLATED LEARNING: SIGNIFICANCE OF DIFFERENCES BETWEEN SIMILAR AND OPPOSED MATERIALS IN TERMS OF *t*-VALUES

	Number Correct Responses		Speed of Responding	
	t	level	*t*	level
1st I.L. Trial				
1 IR	—*	—*	1.29	11%
2 IR	—	—	1.61	6%
4 IR	—	—	1.61	6%
8 IR	1.14	14%	1.57	7%
All Groups	1.12	13%	3.02	1%
I.L. Review Trial				
1 IR	—	—	—	—
2 IR	—	—	—	—
4 IR	—	—	.71	—
8 IR	.62	—	—	—

* Not computed, as differences smaller than those given.

[2] The formula used was

$$t = \frac{\frac{\Sigma d}{n}}{\sqrt{\frac{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}{n-1}}}$$

in which d is the difference in number of correct responses (or percentage of '1' correct responses) between similar and opposed materials for a single *S*, and n is the number of such differences (i.e., the number of *S*s). In the case of the retroactive results, *deviations* from the performance on the O.L. review trial were used instead of raw numbers of correct responses (or raw percentages of '1' correct responses). In determining all significance values a 'single-tail' *t*-table was used, indicating the significance of differences in a given direction.

number of similar and opposed interpolated responses, thus serving as his own control. The formula used takes into account the correlation between similar and opposed values due to inter-subject differences.

(1) *Differential proactive interference on 1st I.L. trial.* In terms of the number of correct responses, there are no significant differences between similar and opposed materials for any group of *Ss* on the 1st I.L. trial where such proactive interference should be greatest. However, as shown in Fig. 1, where differences occur they show more interference for opposed materials. On the other hand, when speed of responding is used as the measure, the opposed condition is found significantly slower for all except the 1 IR group, even this difference approaching the 10 percent level. Further inspection of the *t*-values in Table III shows that when all groups are combined (which is legitimate on the first I.L. trial, since differential degrees of interpolated learning have not been given), differences between similar and opposed materials are significant at the one percent level for speed of responding and at the 13 percent level for number of correct responses. It is apparent that *meaningfully opposed responses suffer greater proactive interference than meaningfully similar responses, the effect being most marked when speed of responding is the measure.*

(2) *Strength of interpolated habits, as indexed by the I.L. review trial.* The purpose of varying the criterion number of correct anticipations on the interpolated learning for the four groups of *Ss* was to insure differential strengths of these interpolated habits. That the procedures were effective in accomplishing this is evident from the negatively accelerated curves obtained on the I.L. review trial, for both number and speed measures, as shown in Fig. 1. The equality between similar and opposed materials was re-established by the conclusion of interpolation; referring to the *t*-values in Table III, *there are no significant differences between similar and opposed materials on either measure by the I.L. review trial.*[3]

II. *Retroactive interference.*—Immediately following the review trial on the interpolated learning, *S* began relearning the original materials, all 14 items being included on every trial. The retroactive results are computed in terms of % RI in order to take into account slight variations in the degree of original learning for 1 IR, 2 IR,

[3] Although there are no *significant* differences on the I.L. review trial, the consistently greater speed for opposed responses for all degrees of interpolation suggests that other than chance factors are involved. Paradoxically, it appears that the final greater speed of making opposed responses is due to their initial slowness. The opposed materials required more trials of practice, on the average, to reach the same criterion, with the exception of the 2 IR group. Since items were dropped from the list when the criterion was reached, this means that similar words tended to be removed earlier and hence could suffer weakening of habit strength through a longer period before being tested on the review trial.

4 IR and 8 IR groups. In determining % RI, the usual formula

$$\% RI = \frac{X - R}{R}(100)$$

was used with one major exception. Whereas R is conventionally the score for the 'rest' or so-called 'control' group,[4] in the present case R *is the score on the review trial of the original learning*, no 'rest' condition having been employed. X is the score on a given relearning trial.

(1) *% RI as measured by number of correct responses.* Table IV presents the average % RI values for the five relearning trials for the four degrees of interpolated learning, based on the number of correct responses made. Results for similar and opposed materials are segre-

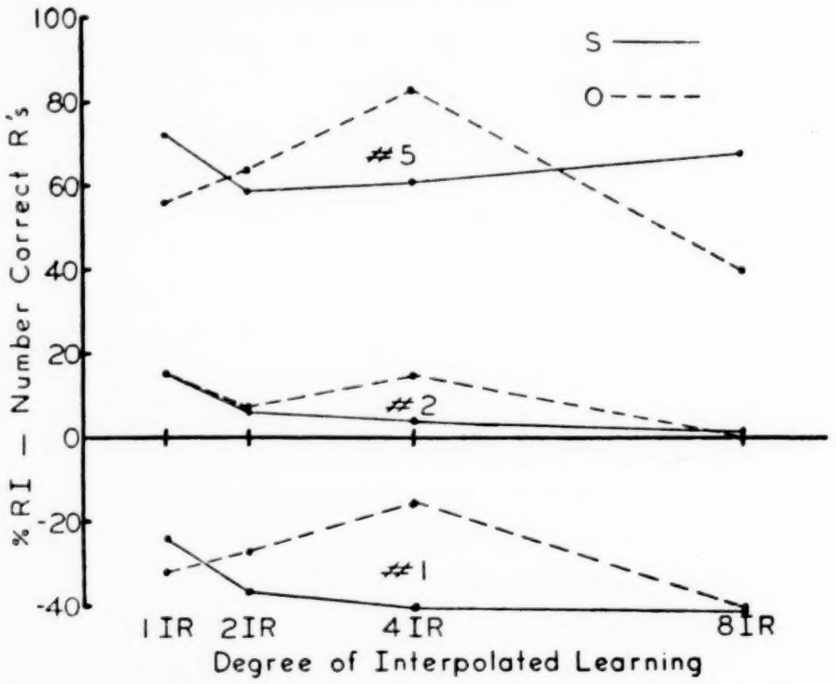

FIG. 2. Relearning: % RI based on mean number of correct responses for relearning trials Nos. 1, 2 and 5

gated. The lower portion of the table gives the significance of the differences between similar and opposed materials in terms of t-values and their levels of significance. Fig. 2 provides graphic representation for the results on trials Nos. 1, 2 and 5.

[4] This conventional method reflects a stage in theory when time itself was postulated as a cause of forgetting, rather than processes occurring in time. That 'rest' conditions are actually a different type of 'work' condition is shown by the fact that they are quite regularly followed by decrements in recall (cf., Underwood, 20), except under special conditions yielding reminiscence. The practical justification, that results of various experiments are rendered comparable, implies that the interference effects of the work during 'rest' will be constant, which appears most unlikely since the character of the 'rest' activity depends entirely upon the whims of the E, some requiring their Ss to read *College Humor*, others having them learn color-names, and still others having them work on jigsaw puzzles. The method employed here has the added advantage that each S serves as his own control, making possible the use of statistical procedures which take into account inter-subject variability.

TABLE IV

RELEARNING OF O.L. AS MEASURED BY NUMBER OF CORRECT RESPONSES: % RI VALUES AND SIGNIFICANCE OF DIFFERENCES FOR SIMILAR AND OPPOSED MATERIALS

	% RI: Relearning Trials									
	No. 1		No. 2		No. 3		No. 4		No. 5	
	S	O	S	O	S	O	S	O	S	O
1 IR	−24	−32	15	15	39	27	43	39	72	56
2 IR	−36	−27	6	7	27	25	46	55	59	64
4 IR	−40	−15	4	15	33	50	61	59	61	83
8 IR	−41	−40	2	0	32	8	58	34	68	40

	Significance of Differences: Relearning Trials									
	No. 1		No. 2		No. 3		No. 4		No. 5	
	t	level	*t*	level	*t*	level	*t*	level	*t*	level
1 IR	.91	—	.00	—	.25	—	.20	—	.67	—
2 IR	.47	—	.00	—	.55	—	.67	—	.52	—
4 IR	2.08	3%	.35	—	1.00	—	.43	—	1.48	8%
8 IR	.78	—	.10	—	1.07	—	2.18	2%	1.67	6%

When number of correct responses is the measure, the similar condition yields typical (**9, 10, 18**) retroactive results: % RI is a negatively accelerated increasing function of the amount of interpolated learning, and recovery is somewhat more rapid with high degrees of interpolation. Opposed materials give a very different pattern, mainly due to the relative superiority of the 4 IR group. Since this group is somewhat atypical throughout the data, it is doubtful if much weight should be given to this anomalous trend. Turning to the differences between similar and opposed conditions, *on the 1st relearning trial, with the exception of the 4 IR group, there are no significant differences in amount of retroactive interference as measured by the number of correct responses.* With the highest degree of interpolation, 8 IR, opposed materials yield more interference than similar materials, differences being significant at about the five percent level on trials Nos. 4 and 5. For both types of material retroactive interference is limited to the first (recall) trial, relative facilitation being found on trial No. 2 and beyond.

(2) *% RI as measured by speed of responding.* Table V and Fig. 3 present equivalent data using speed of responding as a measure. The opposed condition yields typical retroactive results, as evidenced by the negatively accelerated curve for the 1st relearning trial in Fig. 3. For all degrees of interpolation, interference for opposed materials persists through the second relearning trial. Conversely, the major

effect of interpolating similar materials on the speed of responding is *facilitative*. Only the 2 IR group shows a speed decrement from the O.L. review trial level, and the degree of facilitation tends to be greater with higher amounts of interpolation on all relearning trials.

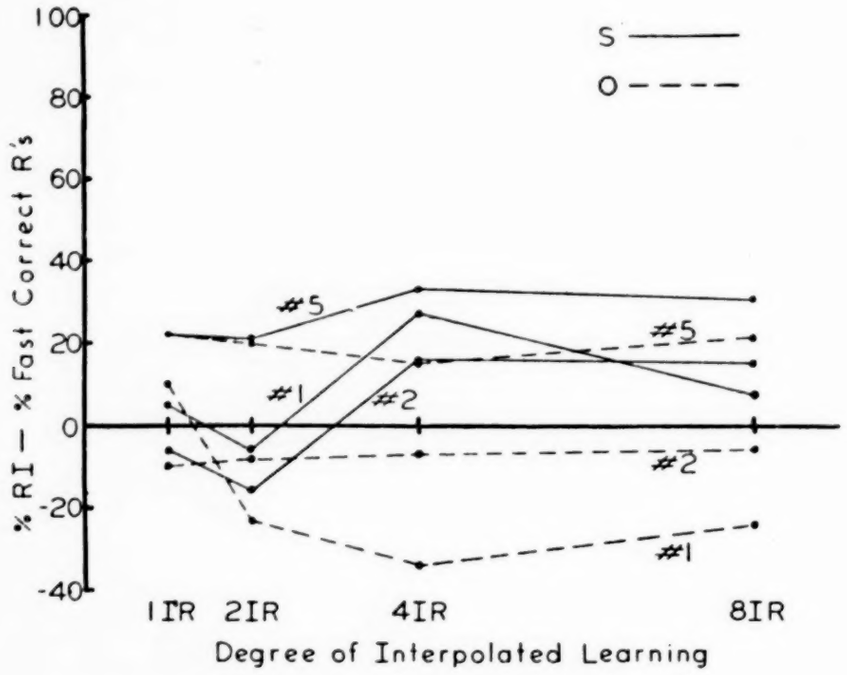

FIG. 3. Relearning: % RI based on percent fast correct responses for relearning trials Nos. 1, 2 and 5

The fact that trial No. 2 shows greater interference than trial No. 1 for the similar condition is difficult to explain; it may indicate that self-produced aids in discriminating between the lists were dropping out. (*S*s were told, at the end of interpolation, that they were going back to the first list.)

Turning finally to the significance of differences between similar and opposed conditions, as evaluted in the table of *t*-values (lower portion of Table V), it may be seen that on the 1st relearning trial, where retroactive effects should be maximal, all degrees of interpolation except the lowest, 1 IR, show similar materials to produce significantly less interference than opposed materials, the *P*-values corresponding to the given *t*'s being .97, .92 and .92 for groups 2 IR, 4 IR and 8 IR.[5] Although differences on trials other than the first

[5] It is likely that these estimates of significance for speed of responding are underestimated. In order to utilize a statistical analysis which takes into account inter-subject variance, it was necessary to compute % '1' correct responses for individual *S*s. The total possible number of correct responses in each meaningful condition for each *S* was only seven, and many speed percentages were based on totals as small as one or two. This results in extremely high variability among the values from which *t*'s were obtained, as well as slight changes in mean values.

TABLE V

RELEARNING OF O.L. AS MEASURED BY SPEED OF RESPONDING: % RI VALUES AND SIGNIFICANCE OF DIFFERENCES FOR SIMILAR AND OPPOSED MATERIALS

	% RI: Relearning Trials									
	No. 1		No. 2		No. 3		No. 4		No. 5	
	S	O	S	O	S	O	S	O	S	O
1 IR	5	10	−6	−10	2	7	10	13	22	22
2 IR	−6	−23	−16	−8	0	8	16	23	21	20
4 IR	27	−34	16	−7	14	−2	18	3	33	15
8 IR	7	−24	15	−6	11	7	33	19	30	21

	Significance of Differences: Relearning Trials									
	No. 1		No. 2		No. 3		No. 4		No. 5	
	t	level	*t*	level	*t*	level	*t*	level	*t*	level
1 IR	.13	—	.61	—	.68	—	.70	—	.22	—
2 IR	1.94	3%	.28	—	.41	—	.04	—	.91	—
4 IR	1.44	8%	.76	—	.20	—	.06	—	.11	—
8 IR	1.43	8%	.61	—	.32	—	.00	—	.31	—

are not significant, with high degrees of interpolation (4 and 8 IR) similar materials show less interference than opposed materials throughout all five relearning trials.

When speed of responding is the measure, the interpolation of opposed materials leads to significantly greater retroactive interference than the interpolation of similar materials. Increasing the amount of interpolated learning affects the two types of material in opposite ways, tending to produce greater interference with opposed material and greater facilitation with similar material.

III. *Analysis of types of errors made.*—Whenever the correct response is *not* made, some indication of the reason for this failure is available. The response may simply not occur, in which case a *blank* is recorded; a word may be intruded from within the list being practiced, *intralist intrusion;* a word may be intruded from outside the experimental materials entirely, *extralist intrusion;* some word from the list previously practiced may be intruded, and this may be either an *unrelated interlist intrusion* (correct and intruding responses having been associated with different stimulus members) or a *related interlist intrusion* (correct and intruding responses having been associated with the same stimulus member). The frequencies of these various types of errors serve as *indexes* of the theoretical causes of interference, and the implications of the following findings for theory will be elaborated in the discussion section. Due to limitation in

space, frequencies of blanks and intrusions under the various conditions are not completely tabularized.

(1) *Errors made during interpolation.* On neither the first I.L. trial nor the I.L. review trial are differences in types of errors of any magnitude found between similar and opposed conditions. On the I.L. review trial frequencies of all types of errors are found to decrease progressively as the amount of interpolated learning is increased, merely reflecting the fact that more correct responses are being made as the lists become better learned. *Related interlist intrusions* (for all interpolated trials combined) *do not increase regularly with the degree of interpolation and they account for only two percent of all errors in each interpolation group.* Since such intrusions are a direct index of the unlearning factor, it would appear that this factor plays a very minor role in this experiment.

(2) *Errors made during relearning.* This experiment also reveals the commonly reported phenomena which lead Melton and Irwin (9) to conclude that some factor other than competition at the point of recall must be operative. *Failures of response (blanks) increase while interlist intrusions decrease, as the amount of interpolated learning is increased,* these effects being most marked on the 1st trial where retroactive interference is maximal. Furthermore, related interlist intrusions occur most frequently in the similar condition.

(3) *Effect of interpolation upon errors made on original materials.* Performance on the *O.L. review trial* provides a convenient evaluation of the pattern of errors being made at the conclusion of the original learning. After the interpolated activity, the Ss return to O.L., and the performance on the *1st relearning trial* may be compared with that on the O.L. review trial to demonstrate the effect of interpolation upon the types of errors made. Since no interlist intrusions appear during the original learning and since extralist intrusions are of insufficient frequency for analysis, only *intralist intrusions* and *blanks* may be compared in this manner.

Table VI presents the relevant data. Percentages that intralist intrusions and blanks are of all errors made on these trials are also included. *From the last (review) trial of the original learning to the first (recall) relearning trial on the same materials, intralist intrusions are reduced by approximately half while failures of response (blanks) increase markedly.* Although an increase in failures of response would be expected to result from interpolation, the large decrease in intralist intrusions is a surprising and hitherto unreported phenomenon (frequencies of intrusions during original learning are not usually measured).

(4) *Distribution of errors through the temporal course of relearning.* While raw frequencies of errors can be used for comparisons among

TABLE VI

RAW FREQUENCIES AND PERCENTAGES OF INTRALIST INTRUSIONS AND FAILURES OF RESPONSE ON O.L. REVIEW TRIAL COMPARED WITH FIRST RELEARNING TRIAL

	1 IR		2 IR		4 IR		8 IR	
	N	%	*N*	%	*N*	%	*N*	%
Intralist								
O.L. Review	72	.43	68	.45	51	.35	45	.33
1st Relearning	36	.18	33	.17	35	.19	34	.17
Blanks								
O.L. Review	84	.50	81	.54	84	.57	89	.64
1st Relearning	112	.55	128	.65	119	.64	145	.73

interpolation groups, they cannot be used to study the temporal course of interference during relearning. As relearning proceeds, more and more correct responses are made and the absolute frequencies of all types of errors necessarily decrease. The *percentage* that each type of error is of all errors made (i.e., the relative importance of each type of error) may be used, and such a procedure is feasible on the retroactive test, since all items were included on each trial.

When the error data are handled in this manner, several theoretically relevant facts appear. *When all data are combined, for both types of meaningful material and for all degrees of interpolation, blanks are found to decrease rapidly in proportional occurrence as relearning precedes, interlist intrusions remain roughly constant, and intralist intrusions increase markedly in relative frequency.* These facts are shown in Table VII. It should be borne in mind that the total fre-

TABLE VII

PERCENTAGE THAT EACH TYPE OF ERROR IS OF ALL ERRORS MADE ON EACH RELEARNING TRIAL

(All Degrees of Interpolation Combined)

	No. 1	No. 2	No. 3	No. 4	No. 5
Blanks	.65	.56	.49	.46	.46
Intralist Intrusions	.17	.27	.33	.36	.40
Extralist Intrusions	.03	.04	.03	.03	.02
Unrelated Interlist Intrusions	.08	.05	.06	.07	.05
Related Interlist Intrusions	.07	.08	.09	.08	.07

quency of all errors is diminishing during relearning (actually from 785 errors on trial No. 1 to only 285 on trial No. 5). These findings indicate that errors attributable to interlist effects (intrusions and blanks based on interlist intrusion tendency) diminish rapidly during

relearning, leaving within-list confusion as a more persistent cause of interference. This finding relates directly to the oft-repeated observation that retroactive interference, when measured in number of correct responses, is of very short duration.

When the temporal course of errors during relearning is broken down in terms of degrees of interpolation, significant differences appear between similar and opposed materials. Fig. 4 compares the mean percent of all errors that are blanks on the *first* two relearning trials with the mean percent of blanks on the *last* two relearning trials. Although for both types of material the all-over percentages of blanks increase as a function of the amount of interpolated learning, the *locus* of the effect is entirely different. *With similar materials,*

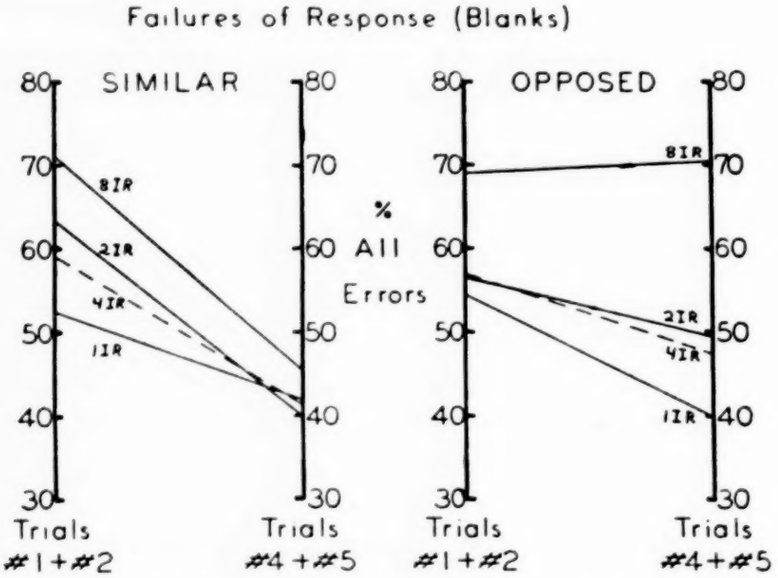

FIG. 4. Percentages that failures of response (blanks) are of all errors made on relearning trials Nos. 1 and 2 compared with relearning trials Nos. 4 and 5, similar and opposed conditions contrasted

failures of response (blanks) appear more in the early trials but then diminish rapidly as the degree of interpolation is increased from 1 IR to 8 IR; conversely, with opposed materials, failures of response as a cause of interference persist longer through the relearning trials as the degree of interpolation is increased. The results for the 4 IR group are again somewhat out of line.

Fig. 5 provides an equivalent analysis of interlist intrusions. *Unrelated* interlist intrusions (upper figures), words previously associated with different stimuli than the correct responses, display no differences between similar and opposed conditions, in both cases decreasing in proportional occurrence as the amount of interpolated learning increases. Analysis of *related* interlist intrusions (lower figures), those having been associated with the same stimulus mem-

bers as the correct responses, does reveal a clear-cut difference. For the opposed condition, with the exception of the 4 IR group, percentage of these intrusions varies with neither the degree of interpolation nor with the course of relearning. On the other hand, *for the similar condition the locus of maximum related interlist intrusion shifts consistently toward later relearning trials as the degree of interpolated learning is increased.* Whereas with the lowest degree of interpolation the maximum is at trial No. 1, with the highest degree of interpolation the maximum is at trial No. 5.

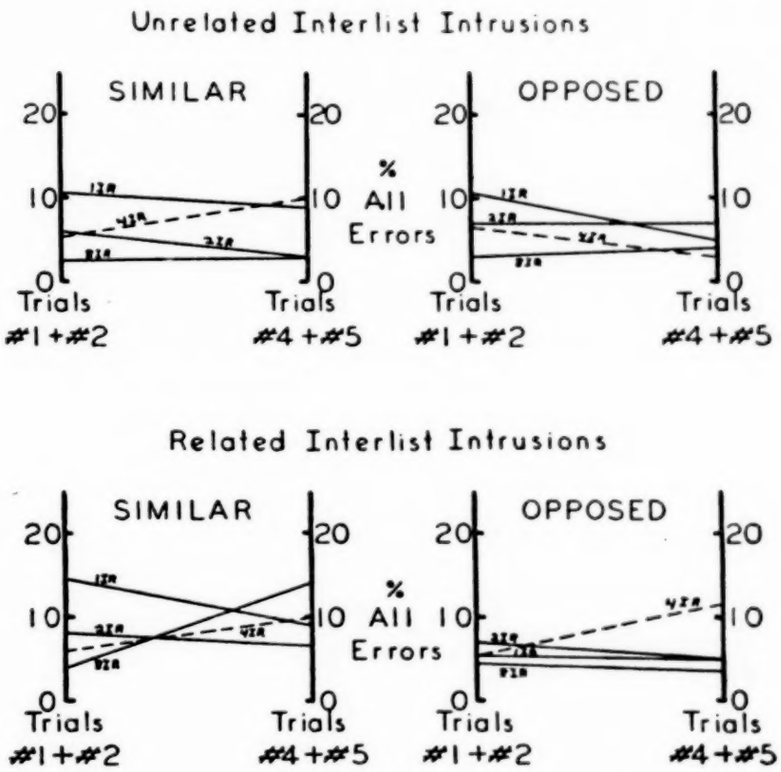

FIG. 5. Percentages that unrelated (upper figures) and related (lower figures) interlist intrusions are of all errors made on relearning trials Nos. 1 and 2 compared with relearning trials Nos. 4 and 5, similar and opposed materials contrasted

DISCUSSION

I. *Meaningful similarity and interference in learning.*—The results of this study substantiate those of an earlier investigation (**12**) in showing that when the strict A—B; A—K; A—B paradigm is followed with meaningful materials, similar interpolation produces less proactive and retroactive interference than opposed interpolation. It extends the results of the previous study by showing that the greater decrement for opposed materials is attributable largely to an increase in the latencies with which responses are made, rather than to a reduction in the number of correct responses. This finding contradicts

the widely encounted generalization that "the greater the meaningful similarity among successively practiced materials, the greater the retroactive interference." This statement is based mainly upon the experimental studies of McGeoch and his associates (**3**, **6**, **7**), all employing a paradigm in which *both* stimulus and response members varied; given these conditions of learning, the results obtained are interpretable by the same principles followed here (cf., Osgood, **12**). To the extent to which competition and two-factor theories must predict increased interference with increasing meaningful similarity, they are found inadequate.

II. *Evaluation of the two-factor theory.*—The two-factor theory attributes total RI to (a) competition of original and interpolated responses at the point of recall and (b) unlearning of the original responses during interpolation. Since, as Underwood (**20**) has pointed out, the effect of unlearning is to reduce the effectiveness with which the original responses compete at the point of recall, total RI is still attributable to interference of the interpolated materials during the attempted recall of the original materials.

(1) *Competition factor.* In terms of such a theory, failures of response (blanks) are due to implicit intrusions made in the response competition situation. However, in this experiment as well as others, the frequency of overt interlist intrusions (index of total interlist intrusion tendency) is not found to correlate with the frequency of blanks. Overt interlist intrusions decrease in frequency while blanks increase, as the degree of interpolated learning is increased.

(2) *Unlearning factor.* To account for this discrepancy, Melton and Irwin (**9**) postulated that unlearning of the original responses occurs during interpolation. But, in exactly the same sense that frequency of overt interlist intrusions during relearning is an index of the strength of the competition factor, *so the frequency of overt intrusions of related* [6] *original responses during interpolation is an index of the amount of unlearning.* That is, in order for an original response to be 'punished' and unlearned, it must occur either overtly or implicitly. The data from this experiment show that the actual frequency of such intrusions is infinitesimally small. In the group where the number of these intrusions was the greatest, there were only 20 cases on all interpolated trials combined; this means an average frequency of one case of an 'unlearning' intrusion on only one of the 14 items and on only a single trial, for each *S*.

[6] It is theoretically unjustifiable in such an analysis to lump together all interlist intrusions. If during interpolation an *S* makes an *unrelated* interlist intrusion (i.e., of a word associated with a stimulus different from the correct response), this is an adequate condition for further discrimination among the original list items—theoretically, the *S* is learning not to make within-list errors on the original list.

(3) *Underwood's revision of the two-factor theory.* Underwood (**20**) has elaborated the two-factor theory so as to incorporate certain of these phenomena. The amount and character of the total interference is postulated to be a function of two variables: (a) the strength of the interpolated responses, this being a negatively accelerated function of the degree of interpolated learning, and (b) the degree of differentiation between the competing response systems, this being a direct function of the degree of learning of the lists and an inverse function of the time interval between the end of interpolation and the beginning of relearning. In terms of these constructs, the increase in blanks and the concomitant decrease in interlist intrusions (as the degree of interpolated learning is increased) are due to improving differentiation among original and interpolated responses. This improved differentiation leads to recognition of interpolated responses as such and hence their rejection; overt intrusions are inhibited in this manner, but the process requires time and therefore more failures of response occur.

When Underwood's hypotheses are tested against the data of this experiment, certain difficulties arise. (a) Increasing differentiation between lists as a function of varying the amount of interpolation should be reflected in the behavior of the *S*s; specifically, interlist errors should occur more frequently during interpolation as a basis for learning these discriminations. As has already been pointed out, the frequencies of such intrusions are very small and their relation to degree of interpolation is irregular. (b) Beyond the difficulties in a theory which attributes effects to 'time' itself, in the present study the time interval between the end of interpolation and the relearning of the original list was constant and hence could not have been responsible for the observed differences in frequencies of blanks and interlist intrusions. (c) Finally, although associative strengths for both original and interpolated learning were equated and time intervals were constant, significant differences between similar and opposed materials were obtained. Neither the original form of the two-factor theory nor Underwood's revision of it is capable of incorporating these differences.

III. *Evidence for an inhibitory factor.*—Despite the conventional use of the term 'retroactive *inhibition*,' neither competition nor two-factor theories actually utilize any conception of an inhibitory psychological process. They assume that the response of B to stimulus A fails to occur because another response, *K*, takes its place, overtly or implicitly. *The results of the present experiment are interpreted as indicating the existence of such an inhibitory factor, the nature of which is revealed in the characteristics of the successive learning of meaningfully opposed materials:* (1) increased latency of response in both proactive

and retroactive situations; (2) persistence of blanks through the course of relearning; (3) both latency and persistence of blanks varying directly with the degree of interpolated learning.

Since it has been amply demonstrated that response latency varies inversely with habit strength (Hull, **2**), it appears that a selective weakening of habit strengths occurs in the opposed meaningful condition but not in the similar condition. This weakening of opposed habits has the further characteristic that its amount is a negatively accelerated function of the degree of interpolated learning, implying that the inhibitory process itself may be a habit phenomenon.

An hypothesis of *reciprocal inhibition of antagonistic reactions* has been proposed (**12**) which would generate such phenomena as have been observed. The hypothesis may be stated as follows: *Simultaneous with every increment in excitatory habit tendency in the association of a given stimulus with a given reaction, there is also generated an equal increment of inhibitory habit tendency in the association of the same stimulus with the directly antagonistic reaction.*[7] In other words, simultaneous with learning any response the *S* is also learning not to make the directly antagonistic response. When he learns to clench his fist to a certain stimulus, he is simultaneously learning not to open his hand to that stimulus; when he learns to call Negroes 'bestial,' he is simultaneously learning not to call them 'noble.' This hypothesis originated in work on the function of meaningful opposites in synesthesia (**4**) and social thinking (**15, 16**).

In utilizing this hypothesis here it is assumed that common verbal opposites are mediated by reciprocally antagonistic implicit reactions, and further, that other principles of learning (Hull, **2**) apply to these inhibitory habits as well. Thus, such inhibitory tendencies are assumed to *generalize*, to increase as a *negatively accelerated growth function* of the number of repetitions, and to *summate algebraicly* with the excitatory tendencies associating the same stimulus and reaction (i.e., the inhibitory tendencies set up during interpolation subtracting directly from the excitatory tendencies developed during original learning).

The obtained results of the experiment may now be incorporated. (1) Since the effect of the previous learning has been to reciprocally weaken opposed habits and since latency varies inversely with habit strength, opposed materials show less speed of response than similar materials, for both proactive and retroactive situations. That no significant difference appears when the *number* of correct responses is

[7] A major difficulty with this principle lies in the definition of 'antagonistic.' The provisonial definition adopted here is that one reaction is antagonistic to another when the same effector elements are involved in both, but reciprocally—those elements being excited in one being inhibited in the other and vice versa.

measured reflects the well-known fact that correct reactions may be made when habits are relatively weak, latencies continuing to decrease with further learning. (2) Similar responses show little or no interference in terms of speed of responding, indicating that they have not suffered weakening of habit strength. The fact that speed *facilitation* is found with high degrees of interpolation, presumably indicates generalization of excitatory tendency among similar responses. (3) Since some (not all) failures of response are due to weakened habit strength, blanks persist longer through relearning for opposed materials than for similar materials. (4) Since the strength of inhibitory tendencies is a negatively accelerated function of the number of repetitions, retroactive interference for opposed materials increases with the degree of interpolated learning, according to a negatively accelerated function.

IV. *Interlist intrusions and the competition factor.*—The existence of interlist intrusions during the relearning process is ample evidence that competition among response systems occurs, and interference for similar materials is attributable mainly to this factor. Frequencies of *related* interlist intrusions are the most direct index of such competition, since we are concerned with the comparative tendencies of the same stimulus to evoke two different responses, the original word and the interpolated word.

This is a straightforward discrimination situation: If the original response is much more strongly associated with the stimulus, it will tend to occur as the correct reaction; on the other hand, if the intrusive similar response is stronger, it will tend to occur. However, the incorrect response will frequently be recognized and rejected, resulting in a blank being recorded. From this it follows that *overt* related interlist intrusions in the similar condition should be most frequent when the strengths of the competing habits are approximately equal. This point of equality during the relearning trials will be progressively delayed as the amount of interpolation is increased (i.e., the stronger the intrusive interpolated habits, the more repetitions of the correct responses required to reach equality). It would therefore be predicted that the locus of maximum related interlist intrusions in the similar condition would shift from relearning trial No. 1 toward trial No. 5 as the degree of interpolation is increased. This was the result obtained.

For the opposed meaningful condition, the only situation theoretically adequate to produce overt related interlist intrusions is where the intrusive reaction is much stronger than the correct reaction, and there is no recognition (equality of antagonistic reaction tendencies should result in cancelation and consequent failure of response). It would therefore be anticipated that (a) there should be

a smaller total frequency of such intrusions in the opposed condition than the similar, (b) they should occur chiefly in the early trials on relearning, and (c) their frequency should increase with the degree of interpolated learning. Only the first of these expectations is clearly verified, but the frequencies involved are very small.

V. *The reciprocal relation between interlist and intralist intrusions.*—It has usually been assumed (**5**) that *within-list* interference is unaffected by the interpolated activities, the measured differences in retention between 'work' and 'rest' conditions being attributed to additional interlist effects. A previously unreported phenomenon was observed in this experiment, however: frequencies of within-list intrusions diminish sharply following interpolated activity, and the data as a whole indicate that intrusions from the interpolated list have been substituted for them. That there is this reciprocal relation between these two sources of intrusion is suggested by two facts; (a) where the frequencies of interlist intrusions are greatest (with low degrees of interpolation) within-list intrusions show the greatest diminution, and (b) during the first three relearning trials, the rapid decrease in interlist intrusions is paralleled by an actual increase in the frequencies of within-list intrusions. This finding poses a problem: Why should any retroactive decrement be obtained if the effect of interpolation is merely to substitute between-list errors for within-list errors?

The answer, judging from the results of the present experiment, would appear to be that between-list errors *more* than replace within-list errors, but are largely reflected in implicit rather than overt intrusions. Retroactive decrement (in terms of number of correct responses) was limited to the first recall trial and was attributable to a marked increase in blanks, total overt intrusions not increasing. Following Underwood (**20**), implicit intrusions are assumed to occur when intrusive responses are stronger than correct responses but are *recognized* as belonging on the interpolated list. It would be difficult to explain how such 'recognition' could occur if the A—B; A—K; A—B paradigm demanded in theory was actually met in practice (i.e., if the original and interpolated stimuli were actually identical). However, it is apparent that the *S*s are well aware of which list they are practicing, and the process of recognizing and rejecting wrong responses is evidenced in their introspective reports. The traditional method of successively practicing lists of materials, which has been used in nearly all studies in this field, including this one, does not completely satisfy the requirements of this paradigm.

Finally, it should be noted that the marked reduction in within-list intrusions on the original list following interpolation may represent a valid phenomenon independent of the substitution of inter-list errors.

Discriminations built up among the stimuli on the interpolated list may *transfer* to the stimuli on the other list, the amount of such transferred discrimination presumably increasing with the similarity of stimuli on the two lists and the number of interpolated trials. Such a conception of transfer of discrimination has been utilized elsewhere (**11, 14, 17**). Although the high degree of similarity between stimuli on the two lists used here would facilitate such transfer, the increase in discriminability (reduced within-list intrusions) to be expected with increasing amounts of interpolation was not obtained.

Summary

Sets of 14 paired-associates were learned by the anticipation method in which the stimuli were letter-pairs such as *t.g.* and the responses were common adjectives. The traditional retroactive interference paradigm, A—B; A—K; A—B, was followed, one experimental variable being the meaningful relationship between B and K adjectives (similar or opposed) and the other variable being the degree of interpolated learning. Interference was measured both in terms of number of correct responses and in terms of the speed of responding, and complete record was kept of all intrusions and failures of response. The following results were obtained:

1. As measured by number of correct responses, there were no significant differences between similar and opposed materials in either proactive or retroactive interference, with the exception of one degree of interpolation where the result was of somewhat questionable validity.

2. As measured by speed of responding, similar materials yielded significantly less proactive and retroactive interference than opposed materials, the effect of increasing the degree of interpolation being to produce greater speed facilitation with similar material and greater speed decrement with opposed materials.

Interpretation of these results in conjunction with analysis of the types of errors made leads to the following conclusions: (1) The common statement that "the greater the meaningful similarity, the greater the retroactive interference" does not have general validity; (2) the two-factor theory of retroactive interference is not adequate to cover all the relevant phenomena in the field; (3) an additional factor, here postulated to be reciprocal inhibition of antagonistic reactions, is required.

(Manuscript received March 20, 1947)

References

1. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
2. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
3. Johnson, L. M. Similarity of meaning as a factor in retroactive inhibition. *J. gen. Psychol.*, 1933, **9**, 377–388.
4. Karwoski, T. F., Odbert, H. S., & Osgood, C. E. Studies in synesthetic thinking: II. The role of form in visual responses to music. *J. gen. Psychol.*, 1942, **26**, 199–222.
5. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green and Co., 1942.
6. McGeoch, J. A., & McDonald, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, **43**, 579–588.
7. McGeoch, J. A., & McGeoch, G. O. Studies in retroactive inhibition: X. The influence of similarity of meaning between lists of paired associates. *J. exp. Psychol.*, 1937, **21**, 320–329.
8. McGeoch, J. A., & Underwood, B. J. Tests of the two-factor theory of retroactive inhibition. *J. exp. Psychol.*, 1943, **32**, 1–17.
9. Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 175–203.
10. Melton, A. W., & von Lackum, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
11. Orbison, W. D. (Unpublished manuscript)
12. Osgood, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36**, 277–301.
13. Postman, L. Reaction time as a measure of retroactive inhibition. *Amer. Psychologist*, 1946, **1**, p. 259.
14. Robinson, E. S. *Association theory today.* Century Psychology Series, 1932.
15. Stagner, R., & Osgood, C. E. An experimental analysis of a nationalistic frame of reference. *J. soc. Psychol.*, 1941, **14**, 389–401.
16. Stagner, R., & Osgood, C. E. Impact of war on a nationalistic frame of reference: I. Changes in general approval and qualitative patterning of certain stereotypes. *J. soc. Psychol.*, 1946, **24**, 187–215.
17. Thorndike, E. L. *Human learning.* Century Psychology series, 1931.
18. Thune, L. E., & Underwood, B. J. Retroactive inhibition as a function of degree of interpolated learning. *J. exp. Psychol.*, 1943, **32**, 185–201.
19. Twining, P. E. The relative importance of intervening activity and lapse of time in the production of forgetting. *J. exp. Psychol.*, 1940, **26**, 483–501.
20. Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, No. 3.

A PROBABILITY FORMULATION OF THE HUNTER–SIGLER EFFECT

BY HAROLD SCHLOSBERG

Brown University

In 1940, Hunter and Sigler published a paper entitled 'The Span of Visual Discrimination as a Function of Time and Intensity of Stimulation' (**10**). In their experiments they exposed plates containing varying numbers of black dots on an illuminated white field, and determined the amount of light necessary to yield correct reports on 50 percent of the trials. The measurements were made over a wide range of intensities (.016–13.33 ml.) and of durations (.004–4.0 sec.).[1] Certain of their results may be summarized as follows:

1. To obtain 50 percent incidence of correct reports of one dot, the product of intensity and time of exposure ($I \cdot T$) remains constant up to a critical duration, beyond which intensity is the controlling factor.

2. To reach a 50 percent level of correct responses on plates containing two or more dots, successively larger amounts of light are needed. Thus, if one plots thresholds (50 percent) as a function of intensity and time, he obtains a family of parallel curves, with that for one dot requiring the least light, followed by those for 2, 3, 4, 5 and 6 dots in regular order.

3. Curves for 9–12 dots fall in their appropriate places, but do not remain parallel to the others. They show increasing departure from the $I \cdot T = C$ slope, in the direction of greater dependence on time.

The first point needs little discussion here. Hunter and Sigler showed that it is consistent with the literature on the relationship between intensity and duration in determining visual thresholds. The third point would seem to be related to the familiar span of apprehension. Hunter and Sigler devoted little discussion to this phenomenon, and merely pointed out that one goes beyond the level of a single event when he uses plates of eight or more dots, so that *S* must count, and counting takes more time. We are not concerned here with the nature of this limit; it has been called the span of apprehension, the span of attention, and the span of perception.

The second point we mentioned above is less clearly related to known facts. Why should it take more light to see two black dots than it does to see one? Hunter and Sigler made little progress in searching for an explanation, except to point out that the increased number of black dots decreases the area illuminated, which might have the effect of decreased illumination. This explanation could scarcely account for the observed

[1] The present analysis is based entirely on these measurements. Professor Hunter generously furnished the data in both raw and graphic form.

differences, for the dots were small, compared to the area of the illuminated plate.[2] The findings are of such novelty and significance that they demand a better explanation.

A clue to an explanation can be found in the stimulus situation. The dots were .125 in. in diameter, and were exposed at a distance of 11 feet from *S*, subtending a visual angle of 3.32′. This immediately suggests that an acuity problem is involved.[3] But why should it take more light to see two dots than it does to see one? An answer is not hard to find. It involves (1) the well-established fact that the probability of response to a simple visual stimulus increases as a function of intensity and duration of exposure, and (2) the assumption that when two or more stimuli (dots) are to be seen simultaneously, they act as independent events, whose probabilities of detection must be compounded.

Let us take an $I \cdot T$ value that will permit *S* to see one dot on 50 percent of the trials. This means that five times out of 10 the retinal gradient of stimulation is such that a small area, covered by the dot, is significantly lower than its neighboring field. 'Significantly' simply refers to the fact that there must be some sort of a threshold. The threshold undoubtedly varies from point to point on the retina; indeed, such variation is a fundamental assumption in theories of the visual mechanism (**7**). There may also be variation in the alignment of the dark spot in relation to receptive elements. Variations of the central mechanism (**3, 4**) or even of quanta of light (**8**), add to the possibilities. These variables are the factors underlying the statistical determination of a threshold. For example, Hunter and Sigler base all their later treatment on the 50 percent point, which is merely the value at which the probability of seeing a dot is equal to the probability of not seeing it.

Let us consider the $I \cdot T$ value which gives us 50 percent correct reports on one dot. If our assumptions are correct, when we substitute a two-dot plate, *each dot* should have a probability of .50 of being seen. If they behave as independent events, the probability that *S* will see both dots becomes $.50^2$, or .25. With the same $I \cdot T$ value *S* should be able to report three dots correctly $.50^3$, or .125, and *N* dots, $.50^N$. We can generalize this statement to read:

$$P_N = P_1{}^N,$$

[2] The effect of black figures in reducing the average illumination of the field is probably a complex problem. Fortunately, it need not concern us here, for 10 dots would account for a maximum of three percent of the illuminated area. This is probably within the limits of error, especially at short exposures.

[3] The Hunter-Sigler curve for one dot (**10**, Fig. 8) levels off beyond the critical duration at a log *I* value of −1.75. At a similar intensity Graham and Cook (**6**) found that *S* would discriminate a grid with stripes about 3.5′ wide. A more nearly comparable study is that of Miller and Beck (**13**), which was concerned with the detection of small black dots on an illuminated field. The dots were exposed for 30 sec., but any gain from the long exposure was probably balanced out by the requirement that *S* scan a large field and localize the dot. Their smallest dot (3.7′) was a trifle larger than the Hunter-Sigler dot (3.32′) and required less light (log *I* values of −1.9 vs. −1.75 ml.). A slight extrapolation of Miller and Beck's curve (**3**, Fig. 17) suggests that at a level of −1.75 m′ their *S*s would have seen a 3′ dot. Thus, the Hunter-Sigler results are in line with other acuity studies. Incidentally, it should be noted that neither *S* obtained afterimages of the dots (**10**), which might have complicated the situation.

or the percentage of correct responses to N dots is equal to the Nth power of the percentage of correct responses to one dot, if the amount of light is constant. This formulation can be applied to the Hunter-Sigler data in at least two ways.

I. The Prediction of the Form of Certain Psychophysical Functions

Fig. 1 (modified from Hunter and Sigler, Fig. 2) presents a number of curves, all obtained at the same duration (32 ms.), but each at a

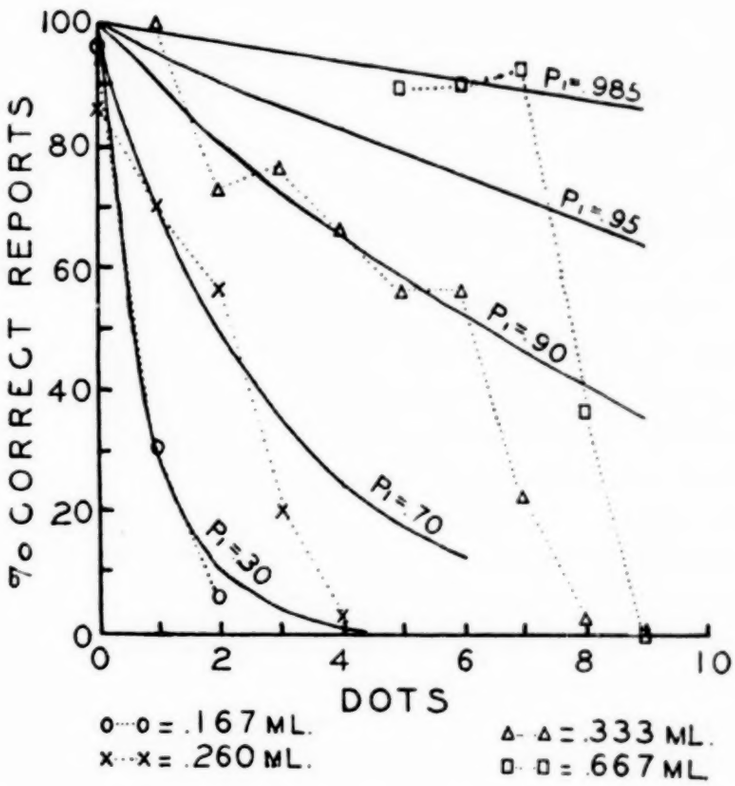

Fig. 1. The percentages of correct reports on the series of 1–9 dot plates, exposed for 32 ms., under four different levels of illumination. There were 30 judgments at each point. The data and dotted lines are from Hunter and Sigler (10, Fig. 2). The solid lines represent expansions of P_1^N, as labelled.

different intensity. The base line indicates the number of dots exposed, and the ordinate gives the percentages of correct identification. Thus, the left hand curve (hollow circles) shows that an intensity of .167 ml. yields 94 percent correct identifications ('no dots') on a blank plate, 30 percent correct reports for one dot, and five percent for two-dot plates. The dotted lines serve merely to connect the points for the reader's convenience; they have no theoretical significance. The solid lines, however, are based on the formulation outlined above. The left hand curve labelled $P_1 = .30$ was chosen to pass through the percentage of correct judgments obtained for *one dot* at .167 ml.

According to our formulation, it passes through the two-dot value at $.30^2$, or .09, and the three-dot value at $.30^3$, or .027. The predicted value for two dots agrees well with the obtained one; no determination was made at this intensity for three dots.

The next curve, $P_1 = .70$, was chosen to pass through the .26 ml. curve (plotted as X) at one dot. It does not give a good fit for the observed values. But the $P_1 = .90$ curve does follow the trend of the .333 ml. values (triangles) as far as six dots. The drop in empirical values for seven and eight dots might be attributed to the limiting effects of the span of perception.[4] This limiting effect is especially clear in the .667 ml. data (hollow squares) at the extreme right. Percentages for 5, 6 and 7 dots suggest a P_1 around .985, but values for 8 and 9 dots fall off rapidly, since they run into the ogive that describes the span of perception (**5**).

The score for our formulation, applied to this set of curves, is one good fit, two possible ones, and a poor one. About the same results turn up if we fit curves to three similar figures in the published paper (**10**). Professor Hunter generously made available the original data, including 22 figures of this type, 11 from each of two highly-practiced *S*s. A family of expanded P_1 curves was drawn on celluloid, and superimposed on each of these figures. Some yielded better fits than those of our Fig. 1, and others were worse, but most of them were about the same. Many of the curves were irregular, for *S*'s task was difficult, and each point involved only 30 determinations. Thus, three judgments would throw a point off by 10 percent. But in spite of such irregularity it was clear that simple compounding of probabilities would account for the higher $I \cdot T$ values required as the number of dots increased. The expanded P_1^N curves obviously gave better fits to the obtained percentages than did the more familiar ogives.

The success of these preliminary studies encouraged an attempt to develop a more precise formulation. To obtain stable empirical values, a number of determinations, each with the same $I \cdot T$ value, were combined; only four durations: 4, 8, 16 and 32 ms., were used, all below the critical duration, beyond which $I \cdot T \neq C$. These combined data are presented in Table I and in Fig. 2. In the figure, each circle represents six sets of 30 judgments each, all obtained at an $I \cdot T$ value (ml. × sec.) between .0050 and .0051. The triangles represent seven sets of data at $I \cdot T$ values between .0098 and .0107. Finally, the crosses are from four sets of data, all obtained at $I \cdot T = .0133$. Thus, each single point represents from 120 to 210 judgments, enough to give stable values.

[4] It will be recalled that our formulation was not intended to apply beyond the limits of a single discriminatory event.

TABLE I

THE AVERAGE PERCENTAGES OF CORRECT JUDGMENTS ON THE SERIES OF ONE- TO NINE-DOT PLATES, AT THREE DIFFERENT LIGHT VALUES

Dots	Quantity of Light (ml. × sec.)		
	.0050–.0051	.0098–.0107	.0133
0	94%	97%	—
1	35%	85%	—
2	12%	70%	90%
3	1%	56%	84%
4	½%	47%	79%
5	—	26%	61%
6		14%	52%
7		2%	46%
8		1%	17%
9		—	2%
N	7 × 30*	6 × 30*	4 × 30*

* Read '7 series of 30 judgments at each point,' etc.

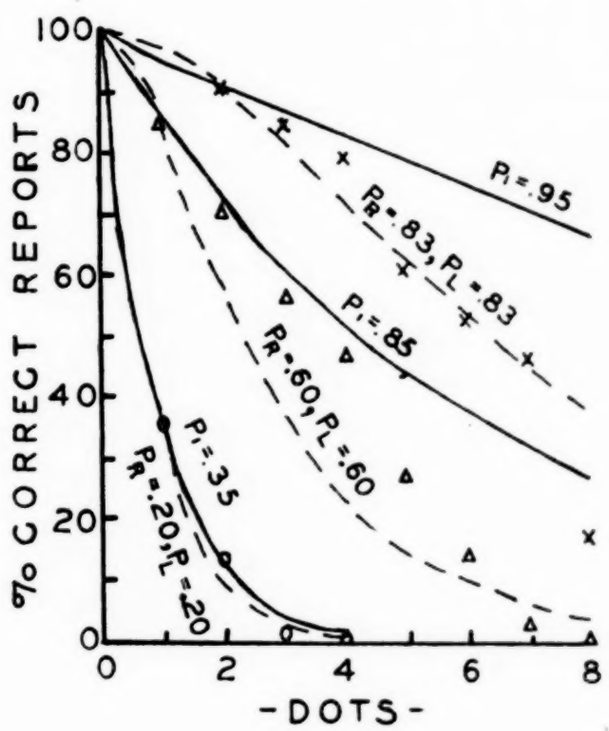

FIG. 2. Average percentages of correct reports on the series of 1–8 dot plates, at light values ($I \cdot T$) given in key. Key: o = .005–.0051 millilambert-sec., Δ = .0098–.0107 millilambert-sec., X = .0133 millilambert-sec. The solid lines represent simultaneously: 1. The expansion of $P_1{}^N$. 2. The expansion of $(1.00 - Q_R Q_L)^N$ (Formulation A, which assumes that *any* dot will be reported if it stimulates *either* eye). The broken lines represent the expansion of $[1.00 - (1.00 - P_R{}^N)(1.00 - P_L{}^N)]$ (Formulation B, which assumes that *all* dots must stimulate the *same* eye if they are to be reported). See text.

The solid lines represent the expanded $P_1{}^N$ percentages. The data for the lowest $I \cdot T$ value fit the formulation nicely, for the $P_1 = .35$ established by reports for one dot yields a curve that passes through the two-dot percentage, and is only two percent too high for three and four dots. The data for the intermediate $I \cdot T$ value fit $P_1 = .85$ very well for two dots. It might be acceptable for three or four dots if it were not clear that the decreasing trend is continued at five, six and seven dots. Similarly, results from the highest intensity give a fair fit for $P_1 = .95$ at three and four dots (in this case the P_1 value was chosen so that the curve would pass through the two dot percentage, since no one-dot plates were exposed). But beyond four dots the observed percentages fall progressively further and further below the predicted ones. One can invoke the span of perception to account for the low observed percentages at eight dots, and perhaps even at seven, but it would hardly account for the low percentages of correct reports on five and six dots.

A second possible distorting factor might be some sort of series effect. An emphasis on accurate observing might be expected to create a tendency to avoid extreme reports. This would account for the observed fact that *S* rarely gave 100 percent correct reports on the blank plates. Similarly, a conservative report on a seven- or eight-dot plate would lead to error, and would depress the percentage of correct responses. On the other hand, it would also be expected to decrease the percentage of correct reports on one-dot plates, resulting in spuriously low P_1 values. Thus, the series effects would be complicated, and would probably neutralize each other. At any rate, it is highly doubtful that they would be large enough to account for the pronounced differences between predicted and obtained percentages on plates with more than five dots.

At this point it became clear that the formulation needed revision. So far, the data had been treated as if they had been obtained in monocular vision; actually they were binocular. There are two limiting ways in which the eyes may act together.

1. One eye may do essentially all of the seeing, either because it is dominant, or because it is more sensitive than the other. In this case our formulation is unchanged.

2. Both eyes may participate equally.[5] In this case the binocular P_1 values are based on a combination of two equal probabilities, P_R, the probability that the right eye will see one dot, and P_L, the probability that the left eye will see it. If we assume that *S* will report

[5] The highly practiced *S*s apparently maintained excellent convergence, using a dimly-lighted fixation ring around the field. It is difficult to predict the effect of slight errors in convergence without monocular data for comparison purposes.

the dot if he sees it with *either* or *both* eyes, we can compute P_R and P_L (in the limiting case, where they are equal) from the formula $P_1 = 1.00 - (Q_RQ_L)$. Then $(P_R + Q_R)(P_L + Q_L)$ when expanded:

$$(1) \quad P_RP_L + P_RQ_L + Q_RP_L + Q_RQ_L$$

gives the appropriate probabilities that the dot will be seen by both eyes (P_RP_L), right eye alone (P_RQ_L), left eye alone (Q_RP_L) and neither eye (Q_RQ_L). For example, let $P_R = P_L = .60$. Both eyes would see the dot $.60 \times .60$, or .36. The right eye would see it $.60 \times .40$, or .24, and the same value would hold for the left eye. Thus, the dot would be seen $.36 + .24 + .24$, or 84 percent of the trials. Hence, if a single dot were seen on 84 percent of the trials, one could consider it the origin of a $P_1 = .84$ curve, according to the original formulation, or as the origin of a $(.60P_R)(.60P_L)$ curve, according to the present formulation.

Now let us extend this formulation to two dots, A and B. Again we must multiply probabilities, for A and B will act as independent events in each eye. Equation (1) above cannot simply be squared, for it is important to keep each term identified by the additional subscripts A and B, to indicate which dot (A or B) and which eye (R or L) are involved in each term. The expansion follows:

$$(2) \quad (P_{RA}P_{LA} + P_{RA}Q_{LA} + Q_{RA}P_{LA} + Q_{RA}Q_{LA})(P_{RB}P_{LB} \cdots Q_{RB}Q_{LB})$$

The first term ($P_{RA}P_{LA}P_{RB}P_{LB}$) will give the probability (in 10,000) that both dots will stimulate both eyes. The second term ($P_{RA}Q_{LA}P_{RB}P_{LB}$) will predict the frequency of cases in which both dots stimulate the right eye, but only dot B stimulates the left eye. The expansion becomes quite laborious, especially when several dots are involved. It is possible to avoid the actual expansion, however, by one of the following methods, depending on the type of binocular interaction assumed.

Formulation A.—If one assumes that any dot will be perceived if it stimulates either eye, the expansion acts exactly as if one were expanding the original P_1 value.[6] It was shown above that $P_1 = .84$ corresponds to ($P_R = .60$, $P_L = .60$) at the one dot value. For two dots $.84^2$, or .7056, is the same value one gets if he expands equation (2) above, with P_R and P_L each equal to .60, and adds all combinations which show that each dot affects at least one eye.

[6] The probability that a single dot will stimulate either or both eyes is equal to $100 - Q_RQ_L$. When N dots are involved, the probability that every dot will stimulate at least one eye becomes $(100 - Q_RQ_L)^N$. But $(100 - Q_RQ_L)$ is equal to P_1, by definition. Hence, the more convenient formulation, P_1^N, can be substituted.

Formulation B.—A different assumption is equally logical: namely, that S will perceive two or more dots *only* if they all stimulate the same eye. The expansion to determine these values is

$$(3) \quad 1.00 - (1.00 - P_R{}^N)(1.00 - P_L{}^N)$$

where N is the number of dots exposed, and P_R and P_L are the (equal) values for right and left eyes, chosen to fit the one dot percentage.

Only the empirical data can decide which of these assumptions is more nearly suitable. The curves predicted under Formulation A are plotted in Fig. 2 in solid lines. They are, of course, the ones already discussed ($P_1 = .95$; $P_1 = .80$; and $P_1 = .35$). The corresponding curves predicted by Formulation B are plotted in broken lines ($P_R = .80$; $P_L = .80$; $P_R = .60$, $P_L = .60$; and $P_R = .20$, $P_L = .20$). It will be seen that the left hand solid and broken curves are almost identical. ($P_1 = .35$) fits the data a trifle better at two dots, but ($P_R = .20$, $P_L = .20$) has a slight superiority at three and four dots. Either curve yields an excellent fit.

The two formulations give quite different curves at other light values. In each case the P_1 or A type of formulation fits better through four dots, but the B type fits better for larger numbers of dots. The break from one curve to the other is not sharp; the data seem to drift from one curve to the other. There is a possible explanation for this. One might expect S to see two dots, even if each dot stimulated only one eye. But it might be much more difficult to integrate a pattern of six dots if they were distributed between the two eyes. In other words, the assumption leading to Formulation A holds when two or three dots are involved, but as we increase the number of dots S tends to behave according to the principle underlying Formulation B.

Discussion.—We started this paper with the question, 'Why did Hunter and Sigler find that it took successively larger amounts of light to recognize 1, 2, 3 $\cdots$ 7 dots?' The answer is simple, if we are content with generalized terms. Hunter and Sigler used the method of constant stimuli, and based all their thresholds on the 50 percent point. But when more than one dot is present on one plate, the dots seem to act as independent events, so that the probabilities are compounded. Thus, the same amount of light that gives 50 percent correct reports on one dot gives $.50^2$ or 25 percent for two dots, and $.50^N$ for N dots. Or put conversely, the amount of light which is required for 50 percent correct reports on two dots will yield $\sqrt{.50}$, or .71 correct reports of one dot. Similarly, the $I \cdot T$ values that give 50 percent correct reports on plates of 3, 4, 5, and 6 dots, will yield, respectively, 79, 84, 87 and 89 percent correct reports on one-dot plates. If

these percentages are plotted against their respective $I \cdot T$ values, one obtains the familiar ogive, which is usually found when percent correct judgments is plotted as a function of intensity in most sense fields. The factors underlying this function, either in general or in the present specific case, need not concern us here. Even the precise form of the function is not important; if we assume that the percentage of correct reports of one dot is a direct function of $I \cdot T$, it follows from the probability formulation that the larger the number of dots exposed on a given plate, the more light will be needed to obtain 50 percent correct reports. This seems to be a perfectly adequate answer to our original question.

But when we applied this formulation to the actual psychophysical functions obtained in the experiment, we found that it was too simple. It had not been intended to fit beyond 6–8 dots, where the limitations imposed by the span of perception show up. But the fit actually became poor even before this limit was reached. Thus, a more precise formulation was necessary. Such a formulation had to take into account the fact that the observations were binocular, rather than monocular.

There have been several attempts to formulate the relationship between monocular and binocular thresholds, considering absolute thresholds (2, 4, 12, 15), difference thresholds (4, 9), and brilliance (1). The formulations vary from those which emphasize simple summation, as of two areas (12), to those which depend entirely on probability (15). One study, that of Hurvich (11), would seem at first glance to be closely related to the present problem. Hurvich exposed plates containing from five to 10 dots, to each of two observers, both monocularly and binocularly. The dots were larger than those used by Hunter and Sigler (15′, compared with 3.32′) which may account for Hurvich's failure to get 'univocal results' in an earlier study. At any rate, in the reported study, Hurvich used reaction time as a measure of sensitivity, and found (1) longer times for monocular than for binocular observations, and (2) an increase in reaction time with increased numbers of dots. The results were interpreted in terms of the Crozier-Holway view, which emphasized the central nervous system, rather than the retina (9). But actually this study offers no help in our formulation. The range studied (5–10 dots) was complicated by the span of perception. Only mean reaction times were given, all obtained at one $I \cdot T$ value. It is difficult to see how one could convert such data to anything useful for our present purpose.[7]

Miller and Beck (13) reported some studies on monocular vs. binocular searching. They found that the increase in relative range brought about by binocular observation fell between that predicted by a probability formu-

[7] The same comment applies to the recent paper by Philip (14). He found a sigmoid relationship between exposure time and correctness of perception of 'color mass' on cards containing 22 or more multi-colored dots.

lation, and one based on simple binocular summation. Their problem was somewhat complicated by a 30-sec. period of observation, during which their Ss searched over a large field for the dot. Their experiment varied from that of Hunter and Sigler in other ways, as in the use of peripheral vision, and in an absence of fixation points.

The only paper which contains monocular and binocular data in a form which is useful to us is that of Pirenne (15). He assumed that the two eyes worked quite independently in seeing a lighted area. Thus, the probability of binocular detection would be the product of the monocular probabilities. His formulation: $P_{bin.} = 1.00 - (1.00 - P_R)(1.00 - P_L)$ is obviously the same as our formulation A (above). Pirenne presented observed percentages for each eye and for binocular vision, and showed that they agreed well with predictions. Although his data were based on peripheral vision, they furnish reasonably strong support for our binocular formulation, as applied to one dot.

But, unfortunately, Pirenne has not considered the problem of several dots, seen binocularly. Indeed, there seem to be no other studies, either monocular or binocular, that are at all comparable with the Hunter-Sigler data in the two to seven dot range. We are thus left with the problem of fitting binocular data over a range of two to seven dots, with no related monocular data, to say nothing of a lack of established theory for monocular vision over this range. If the original Hunter-Sigler apparatus had not been dismantled it would have been fairly simple to run off a series of monocular and binocular determinations at two or three moderate intensities, as .008 and .010 ml. $\times$ sec. We hope that such data can be obtained, either here or elsewhere, to fill this gap.

Lacking these data, we can only support our two formulations on the basis of the way they fit the available results, and in terms of general reasonableness. There would seem to be little doubt that two or three dots can be seen if they stimulate either eye, separately. In other words, Pirenne's probability theory can be extended this far. But as we have indicated above, when the number of dots increases, the two eyes show less and less ability to integrate dots seen by each eye. Thus, six dots must *all* fall on the same eye, or they will not all be seen.

This would seem to conflict with the common observation that two complex monocular fields can be fused if they are not inconsistent. One of the most familiar examples is the stereoscope slide which contains a bird on one half, and a cage on the other. In binocular vision the bird is seen enclosed in the cage. But perhaps these apparently complex objects function as units in perceptual integration. This is clearly the case when we are dealing with them at the level of the span of perception. It is as easy to see five birds as it is to see five or six dots. It is at least possible that something comparable occurs in integrating the fields of the two eyes. Hence, we would expect to see five dots, three stimulating one eye, and two exciting the other, *only* if they were arranged in a pattern, like the face of a die. But if no such pattern exists, we obtain binocular rivalry, and see only the field that receives the larger number of dots, or that is dominant at the time for some other reason.

II. The Prediction of the Frequency of Specific Errors

If a probability formulation is correct, it should be possible to predict errors, for most of the errors will occur because some of the dots on a given plate will not be perceived every time. The attempt to predict the frequency of specific errors was suggested by a minor point raised by Hunter and Sigler. On page 175 (**10**) they presented a count of the number of times five or more dots were reported on plates containing 5, 6, 7 and 8 dots. All plates were exposed at the same $I \cdot T$ value, chosen to give 24 percent reports of five dots on a five-dot plate. On the plates containing 6, 7 and 8 dots, respectively, *S* reported five or more dots with the following frequencies; 43, 64 and 85 percent. If we apply our relatively simple formulation A, we can find the P_1 value that corresponds to 24 percent at five dots. Or, to state it otherwise, solve for P_1 when $P_5 = .24$. The P_1 value is .75. Now, if we expand $(.75P + .25Q)^6$, the first two terms, P^6 and P^5Q, will give us the number of times six and five dots, respectively, will be reported in response to a six-dot plate, under the same light conditions used in determining P_1. Expansions of this type for six-, seven- and eight-dot plates predict the following percentages of reports of five or more dots: six-dot plate, 55 percent; seven-dot plate, 76 percent; eight-dot plate, 89 percent. The predicted values are not very close to the observed ones (43, 64 and 85 percent), but they at least tend in the right direction. Perhaps a more accurate prediction could have been made if we had used the more appropriate formulation B, since it gives a better fit where many dots are involved. But at present it is not obvious how this more elaborate formulation could have been applied in detail. Indeed, it would be easier to do the experiment over again, using monocular vision!

Another way of approaching the same problem is to attempt to predict the specific reports on a single plate. For example, given a plate of five dots, exposed at an $I \cdot T$ value that yields 25 percent correct reports, how frequently should *S* see 0, 1, 2, 3 and 4 dots, respectively? Using the original Hunter-Sigler data sheets, I tallied *all* reports obtained from five-dot plates exposed at intensities which yielded 20–30 percent correct reports. It would have been better to include only those judgments made at one specific $I \cdot T$ value, as I did in Fig. 2, but there would have been too few cases. The selection was well balanced throughout the $P_5 = 20$–30 percent range, so that out of a total of 221 judgments, I obtained 55 (25 percent) reports of five dots. The number of responses of 0, 1, 2, 3, 4 and 5 dots are given in Table II, in the second column. They are converted into percentages in the third column. In the fourth column are the pre-

dicted percentages, based on an expansion of $(.75P + .25Q)^5$, which was chosen to give an approximate fit to the observed percentage of correct responses. The fit seems *too* good, when one remembers that we are using the relatively simple formulation A.

A similar study was made of four-dot plates, exposed at an $I \cdot T$ value which yielded 30–50 percent correct reports. The results are given in the right hand half of Table II. The agreement between observed and predicted reports of 0, 1 and 2 dots is about as good as one would expect from the simple formulation A. It is interesting that Ss reported more dots than were present on three trials. Such errors remind us that there is an organism attached to the eyes and to the occipital cortex!

TABLE II

THE DISTRIBUTION OF REPORTS ON FIVE-DOT PLATES (LEFT) AND ON FOUR DOT PLATES (RIGHT)

R is the number of reports, %R is the same, expressed as %, and P_1 is the value obtained by expanding the given value in the formula $(P + Q)^N$, where *N* is the number of dots exposed. The P_1 values were chosen to give the observed number of correct responses, as indicated by *.

Dots Reported	Five-Dot Plates			Four-Dot Plates		
	R	% R	*P* .75	R	% R	*P* .80
0	3	.01	.00	8	.03	.01
1	5	.02	.02	12	.04	.03
2	19	.09	.09	50	.17	.15
3	60	.27	.26	98	.33	.41
4	79	.36	.40	120	.41	.41*
5	55	.25	.24*	3	.01	.00
Total	221	1.00	1.01	291	.99	1.01

SUMMARY AND CONCLUSIONS

1. Hunter and Sigler found that successively larger amounts $(I \cdot T)$ of light are necessary to yield 50 percent correct reports as the number of dots exposed is increased from one to six or seven. This finding can be understood if we assume that the perception of each dot is an independent event, so that the probability of seeing two dots is less than the probability of seeing either alone.

2. The simplest formulation states that the probability of seeing *N* dots is the *N*th power of the probability of seeing one dot, at the same light value. This formulation gives an approximate fit for the observed psychophysical functions, and even indicates the distribution of false judgments.

3. For a more precise formulation it is necessary to compound the probabilities still further, since the data were obtained with binocular vision. The results suggest that *S* can integrate two or three dots,

even though they stimulate different eyes. But there seems to be a limit to such integration, for five or six dots must *all* stimulate the same eye if they are to be perceived. Finally, at seven or eight dots, *S* reaches another limit, the span of apprehension.

(Manuscript received May 31, 1947)

References

1. Bartley, S. H. *Vision.* New York: Van Nostrand Company, 1941.
2. Crawford, B. H. Ocular interaction in its relation to measurements of brightness threshold. *Proc. Roy. Soc.*, 1939–40, **128B**, 552–559.
3. Crozier, W. J. The theory of the visual threshold. I. Time and intensity. *Proc. nat. Acad. Sci.*, 1940, **26**, 54–60.
4. Crozier, W. J., & Holway, A. H. Theory and measurement of visual mechanisms. II. Threshold stimulus intensity and retinal position. *J. gen. Physiol.*, 1938–9, **22**, 351–364. III. ΔI as a function of area, intensity, and wave-length, for monocular and binocular stimulation. *J. gen. Physiol.*, 1939–40, **23**, 101–141.
5. Fernberger, S. W. A preliminary study of the range of visual apprehension. *Amer. J. Psychol.*, 1921, **32**, 121–133.
6. Graham, C. H., & Cook, C. Visual acuity as a function of intensity and exposure time. *Amer. J. Psychol.*, 1937, **49**, 654–661.
7. Hecht, S. Vision: II. The nature of the photoreceptor process. Chap. 14 in: Murchison, C. (ed.), *A handbook of general experimental psychology.* Worcester, Mass.: Clark Univ. Press, 1934.
8. Hecht, S. Energy and vision. *American Scientist*, 1944, **32**, 159–177.
9. Holway, A. H., & Hurvich, L. M. Visual differential sensitivity and retinal area. *Amer. J. Psychol.*, 1938, **51**, 687–675.
10. Hunter, W. S., & Sigler, M. The span of visual discrimination as a function of time and intensity of stimulation. *J. exp. Psychol.*, 1940, **26**, 160–179.
11. Hurvich, L. M. The range of apprehension and sensory discrimination. *J. exp. Psychol.*, 1940, **27**, 313–317.
12. Lythgoe, R. J., & Phillips, L. R. Binocular summation during dark adaptation. *J. Physiol.*, 1938, **91**, 427–436.
13. Miller, C. W., & Beck, L. H. The effects of night binocular design features on the visibility of targets at low levels of illumination. 1945. OSRD Report No. 6128 (Released, but not otherwise published).
14. Philip, B. R. The relationship of exposure time and accuracy in a perceptual task. *J. exp. Psychol.*, 1947, **37**, 178–186.
15. Pirenne, M. H. Binocular and uniocular threshold of vision. *Nature*, 1943, **152**, 698–699.

'SUPERSTITION' IN THE PIGEON

BY B. F. SKINNER

Indiana University

To say that a reinforcement is contingent upon a response may mean nothing more than that it follows the response. It may follow because of some mechanical connection or because of the mediation of another organism; but conditioning takes place presumably because of the temporal relation only, expressed in terms of the order and proximity of response and reinforcement. Whenever we present a state of affairs which is known to be reinforcing at a given drive, we must suppose that conditioning takes place, even though we have paid no attention to the behavior of the organism in making the presentation. A simple experiment demonstrates this to be the case.

A pigeon is brought to a stable state of hunger by reducing it to 75 percent of its weight when well fed. It is put into an experimental cage for a few minutes each day. A food hopper attached to the cage may be swung into place so that the pigeon can eat from it. A solenoid and a timing relay hold the hopper in place for five sec. at each reinforcement.

If a clock is now arranged to present the food hopper at regular intervals *with no reference whatsoever to the bird's behavior*, operant conditioning usually takes place. In six out of eight cases the resulting responses were so clearly defined that two observers could agree perfectly in counting instances. One bird was conditioned to turn counter-clockwise about the cage, making two or three turns between reinforcements. Another repeatedly thrust its head into one of the upper corners of the cage. A third developed a 'tossing' response, as if placing its head beneath an invisible bar and lifting it repeatedly. Two birds developed a pendulum motion of the head and body, in which the head was extended forward and swung from right to left with a sharp movement followed by a somewhat slower return. The body generally followed the movement and a few steps might be taken when it was extensive. Another bird was conditioned to make incomplete pecking or brushing movements directed toward but not touching the floor. None of these responses appeared in any noticeable strength during adaptation to the cage or until the food hopper was periodically presented. In the remaining two cases, conditioned responses were not clearly marked.

The conditioning process is usually obvious. The bird happens to be executing some response as the hopper appears; as a result it tends to repeat this response. If the interval before the next presentation is not so great that extinction takes place, a second 'contingency' is probable. This strengthens the response still further and subsequent reinforcement becomes more probable. It is true that some responses go unreinforced and

some reinforcements appear when the response has not just been made, but the net result is the development of a considerable state of strength.

With the exception of the counter-clockwise turn, each response was almost always repeated in the same part of the cage, and it generally involved an orientation toward some feature of the cage. The effect of the reinforcement was to condition the bird to respond to some aspect of the environment rather than merely to execute a series of movements. All responses came to be repeated rapidly between reinforcements—typically five or six times in 15 sec.

The effect appears to depend upon the rate of reinforcement. In general, we should expect that the shorter the intervening interval, the speedier and more marked the conditioning. One reason is that the pigeon's behavior becomes more diverse as time passes after reinforcement. A hundred photographs, each taken two sec. after withdrawal of the hopper, would show fairly uniform behavior. The bird would be in the same part of the cage, near the hopper, and probably oriented toward the wall where the hopper has disappeared or turning to one side or the other. A hundred photographs taken after 10 sec., on the other hand, would find the bird in various parts of the cage responding to many different aspects of the environment. The sooner a second reinforcement appears, therefore, the more likely it is that the second reinforced response will be similar to the first, and also that they will both have one of a few standard forms. In the limiting case of a very brief interval the behavior to be expected would be holding the head toward the opening through which the magazine has disappeared.

Another reason for the greater effectiveness of short intervals is that the longer the interval, the greater the number of intervening responses emitted without reinforcement. The resulting extinction cancels the effect of an occasional reinforcement.

According to this interpretation the effective interval will depend upon the rate of conditioning and the rate of extinction, and will therefore vary with the drive and also presumably between species. Fifteen sec. is a very effective interval at the drive level indicated above. One min. is much less so. When a response has once been set up, however, the interval can be lengthened. In one case it was extended to two min., and a high rate of responding was maintained with no sign of weakening. In another case, many hours of responding were observed with an interval of one min. between reinforcements.

In the latter case, the response showed a noticeable drift in topography. It began as a sharp movement of the head from the middle position to the left. This movement became more energetic, and eventually the whole body of the bird turned in the same direction, and a step or two would be taken. After many hours, the stepping response became the predominant feature. The bird made a well defined hopping step from the right to the left foot, meanwhile turning its head and body to the left as before.

When the stepping response became strong, it was possible to obtain a mechanical record by putting the bird on a large tambour directly connected with a small tambour which made a delicate electric contact each time

stepping took place. By watching the bird and listening to the sound of the recorder it was possible to confirm the fact that a fairly authentic record was being made. It was possible for the bird to hear the recorder at each step, but this was, of course, in no way correlated with feeding. The record obtained when the magazine was presented once every min. resembles in every respect the characteristic curve for the pigeon under periodic reinforcement of a standard selected response. A well marked temporal discrimination develops. The bird does not respond immediately after eating, but when 10 or 15 or even 20 sec. have elapsed it begins to respond rapidly and continues until the reinforcement is received.

In this case it was possible to record the 'extinction' of the response when the clock was turned off and the magazine was no longer presented at any time. The bird continued to respond with its characteristic side to

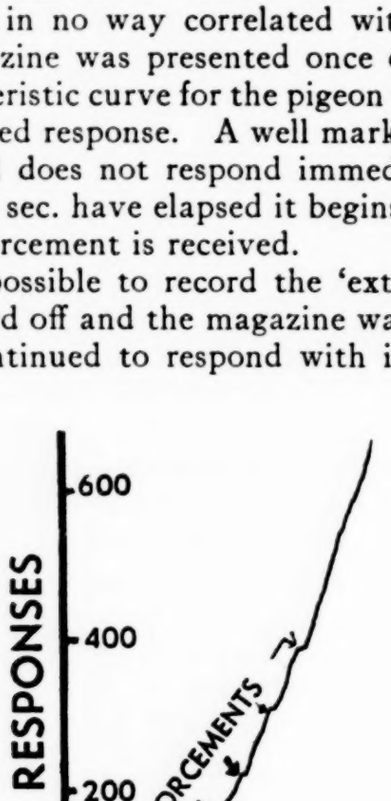

FIG. 1. 'Reconditioning' of a superstitious response after extinction. The response of hopping from right to left had been thoroughly extinguished just before the record was taken. The arrows indicate the automatic presentation of food at one-min. intervals without reference to the pigeon's behavior.

side hop. More than 10,000 responses were recorded before 'extinction' had reached the point at which few if any responses were made during a 10 or 15 min. interval. When the clock was again started, the periodic presentation of the magazine (still without any connection whatsoever with the bird's behavior) brought out a typical curve for reconditioning after periodic reinforcement, shown in Fig. 1. The record had been essentially horizontal for 20 min. prior to the beginning of this curve. The first reinforcement had some slight effect and the second a greater effect. There is a smooth positive acceleration in rate as the bird returns to the rate of responding which prevailed when it was reinforced every min.

When the response was again extinguished and the periodic presentation of food then resumed, a different response was picked up. This consisted of a progressive walking response in which the bird moved about the cage.

The response of hopping from side to side never reappeared and could not, of course, be obtained deliberately without making the reinforcement contingent upon the behavior.

The experiment might be said to demonstrate a sort of superstition. The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking. There are many analogies in human behavior. Rituals for changing one's luck at cards are good examples. A few accidental connections between a ritual and favorable consequences suffice to set up and maintain the behavior in spite of many unreinforced instances. The bowler who has released a ball down the alley but continues to behave as if he were controlling it by twisting and turning his arm and shoulder is another case in point. These behaviors have, of course, no real effect upon one's luck or upon a ball half way down an alley, just as in the present case the food would appear as often if the pigeon did nothing—or, more strictly speaking, did something else.

It is perhaps not quite correct to say that conditioned behavior has been set up without any previously determined contingency whatsoever. We have appealed to a uniform sequence of responses in the behavior of the pigeon to obtain an over-all net contingency. When we arrange a clock to present food every 15 sec., we are in effect basing our reinforcement upon a limited set of responses which frequently occur 15 sec. after reinforcement. When a response has been strengthened (and this may result from one reinforcement), the setting of the clock implies an even more restricted contingency. Something of the same sort is true of the bowler. It is not quite correct to say that there is no connection between his twisting and turning and the course taken by the ball at the far end of the alley. The connection was established before the ball left the bowler's hand, but since both the path of the ball and the behavior of the bowler are determined, some relation survives. The subsequent behavior of the bowler may have no effect upon the ball, but the behavior of the ball has an effect upon the bowler. The contingency, though not perfect, is enough to maintain the behavior in strength. The particular form of the behavior adopted by the bowler is due to induction from responses in which there is actual contact with the ball. It is clearly a movement appropriate to changing the ball's direction. But this does not invalidate the comparison, since we are not concerned with what response is selected but with why it persists in strength. In rituals for changing luck the inductive strengthening of a particular form of behavior is generally absent. The behavior of the pigeon in this experiment is of the latter sort, as the variety of responses obtained from different pigeons indicates. Whether there is any unconditioned

behavior in the pigeon appropriate to a given effect upon the environment is under investigation.

The results throws some light on incidental behavior observed in experiments in which a discriminative stimulus is frequently presented. Such a stimulus has reinforcing value and can set up superstitious behavior. A pigeon will often develop some response such as turning, twisting, pecking near the locus of the discriminative stimulus, flapping its wings, etc. In much of the work to date in this field the interval between presentations of the discriminative stimulus has been one min. and many of these superstitious responses are short-lived. Their appearance as the result of accidental correlations with the presentation of the stimulus is unmistakable.

(Manuscript received June 5, 1947)

DISCRIMINATIVE CONDITIONING. II. EFFECTS OF A PAVLOVIAN CONDITIONED STIMULUS UPON A SUBSEQUENTLY ESTABLISHED OPERANT RESPONSE *

BY WILLIAM K. ESTES

Indiana University

In a previous experiment (1) it was found that a tone which had repeatedly preceded the presentation of food to a group of rats became a discriminative stimulus [1] for a lever-pressing response which had previously been reinforced with food, although the response had never been associated with the tone prior to the test trial. These results were interpreted in terms of the conditioning to the tone of a changed state of the organism which might conveniently be described as 'anticipation' of the reinforcing stimulus. A possible objection to this interpretation is that since the lever-pressing response had been reinforced in the experimental situation before the discriminative training, the tendency to make the response might have been present at some strength during the series of tone-food presentations and therefore directly associated with the tone.

The following experiment has been designed to test this latter interpretation by determining the effect of a tone which repeatedly precedes the presentation of food upon a response which is not added to the organism's repertoire until *after* completion of the series of tone-food presentations. In this experiment, the animals were first given a number of presentations of a tone followed by food; secondly, a lever-pressing response was reinforced with food; and finally, a test was made of the effect of the tone upon the strength of the response.

Subjects and Apparatus

The *S*s were eight albino rats, about 60 days old. The two experimental groups, hereafter referred to as Group A and Group B, each included one male and three females. The animals were fed dry Purina dog food for one hour immediately after running each day. Water was available in the living cages at all times.

The apparatus has been described in detail elsewhere (3). Each rat was enclosed during the daily one-hour experimental period in a light-proof and nearly sound-proof box equipped with a retractable lever and a magazine which could be set to reinforce responses to the lever

* This experiment was conducted in the psychological laboratories of the University of Minnesota.

[1] That is, presentation of the tone during the test period resulted in a significant increase in rate of lever-pressing.

with small (.05 gm.) pellets of food. Under the procedure of periodic reinforcement, a control clock was set to reinforce single responses to the lever every four min., intervening responses going unreinforced. When the levers were present in the boxes, a curve representing number of responses vs. time was recorded for each rat. The discriminative tones were produced by phones attached to a 60 cycle A.C. transformer.

Procedure

The experimental design is summarized in Table I. During the first six days of the experiment the levers were removed from the experimental boxes. On the first day, all eight animals were placed in the boxes and the magazines set to deliver one food pellet per min. Each rat was

TABLE I

Procedure

	Experimental Periods							
	1	2	3	4	5	6	7	8
Group A	S^R	--30(T....S^R)----					R--S^R	T--R
Group B	S^R	-------50(T....S^R)---------					R--S^R	T--R

R = Lever-pressing response; T = Tone; S^R = Operation of food magazine.

returned to the living cage as soon as it had learned to respond promptly to the click of the magazine by approaching the food tray and seizing the food pellet. Beginning with the second day, both groups received discriminative training under the following procedure. The animals were placed in the experimental boxes at the usual time, and after a few minutes the tones were turned on and allowed to sound for 60 sec., then a pellet of food was delivered to each rat by means of the magazine, and the tones were turned off. This combination of tone and food was repeated 10 times during each of three daily experimental periods for Group A and 10 times during each of five periods for Group B. On the seventh day of the experiment, the levers were introduced into the boxes *for the first time* and all eight rats were conditioned to press the lever for food. Each rat was reinforced for the first 10 responses to the lever, then given one hour of periodic reinforcement of the response. On the eighth day, the levers were again present, but the food magazines were disconnected so that all responses went unreinforced. During this extinction period the tone was presented to the rats in Group A for two 10-min. intervals and to the rats in Group B for one 10-min. and one 15-min. interval in order to determine whether it would affect the strength of the lever-pressing response.

Results and Discussion

Mean extinction records for the two groups are presented in Fig. 1 and Table II. For convenience in analysis of the data, the extinction period has been divided into 12 five-min. intervals. The tone was sounding during the seventh, eighth, eleventh, and twelfth of these intervals for Group A and during the fifth, sixth, tenth, eleventh, and twelfth for Group B. Curve C in Fig. 1 is a control extinction curve for a group of four rats taken from a concurrent experiment. This curve illustrates the extent of irregularity ordinarily observed in an uncomplicated extinction curve for a group of this size. It has been shown in earlier experiments (**2, 4**) that the tone has no unconditioned facilitative effect upon the lever-pressing response.

It is clear that the tone exerted a marked discriminative effect upon the pressing response, although it had never previously been associated with the response. In every case, introduction of the tone was followed by a temporary increase in the rate of responding. This tendency is evident in Group A, but is more clearly demonstrated in Group B. In the latter case, the first administration of the tone

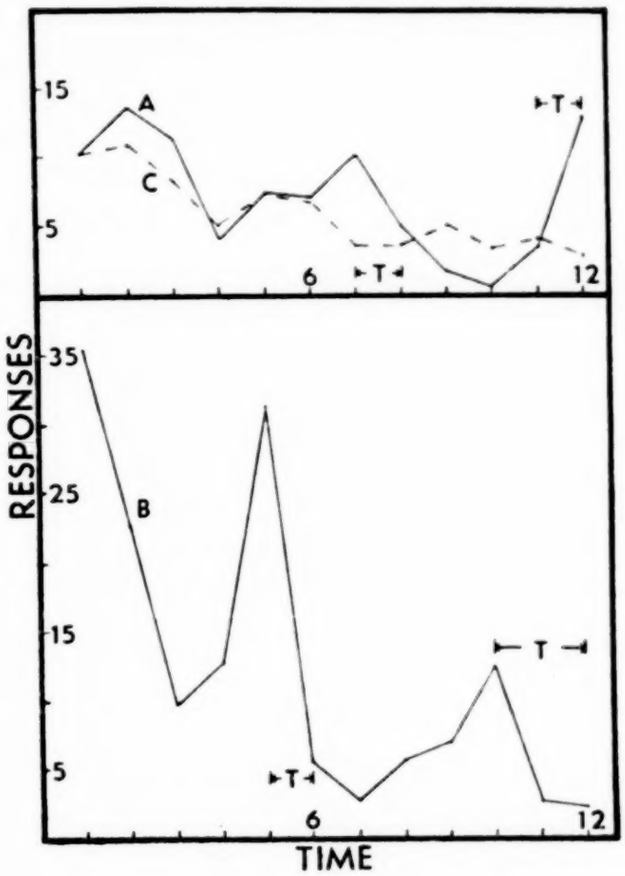

FIG. 1. Mean extinction records plotted by five-min. intervals, showing discriminative effect of the tone upon the rate of responding of Groups A and B. The tone was turned on during intervals 7, 8, 11, and 12 for Group A, and during intervals 5, 6, 10, 11, and 12 for Group B. Curve C is a control extinction curve taken from another experiment.

TABLE II

EXTINCTION RECORDS IN TERMS OF MEAN NUMBER OF RESPONSES PER FIVE-MINUTE INTERVAL

	Interval											
	1	2	3	4	5	6	7	8	9	10	11	12
Group A	10.2	13.5	11.2	4.0	7.2	7.0	10.2*	5.0*	1.5	0.5	3.5*	12.8*
Group B	35.5	22.8	10.0	13.0	31.2*	5.5*	2.8	5.8	7.2	12.5*	2.8*	2.2*
Group C	10.2	10.8	8.2	5.0	7.2	6.8	3.5	3.5	5.0	3.2	4.0	2.8

* Tone sounding during interval

during extinction very nearly restores the rate of responding to the level which obtained at the beginning of the extinction period. The irregularity of the curve for Group A, especially during the first half of the hour, is probably due to the fact that these animals had less opportunity for adaptation of emotional and exploratory reactions to the experimental situation than those of Group B.

These results are very similar to those of the earlier study (1) in which the experimental animals were conditioned to press the lever prior to the series of tone-food presentations. Since in the present experiment the response used to test the discrimination was arbitrarily selected and could never have occurred prior to the tone-food trials, it is reasonable to suppose that lever-pressing is merely one member of an indefinitely large class of responses which could be brought under control of the tone by the same series of operations.

It will be noted that the stimulus relationships in this experiment are similar to those of classical conditioning. In the Pavlovian situation, a tone repeatedly precedes presentation of food to the experimental animal, but only one measure of conditioning, the salivary response to the tone, is obtained. The present experiment indicates that the changed state of the organism which becomes conditioned to the tone in that situation involves generalized discriminative properties which may be measured in terms of the changed strengths of other responses.

Summary

This experiment was designed to determine whether an originally neutral stimulus which has repeatedly preceded the presentation of a reinforcing stimulus (e.g., food) to an organism will exert discriminative control over an operant response subsequently conditioned by means of that reinforcing stimulus. A group of rats was first subjected to a series of presentations of a tone followed by food, then conditioned to press a lever for food reinforcement. Finally, during a test period, the rate of lever-pressing was markedly increased by presentation of the tone and depressed during silent intervals, although the response had not previously been associated with the tone in any way.

It is suggested that responses in an animal's repertoire which have been conditioned by means of the same reinforcing stimulus belong to a functional class, and that an originally neutral stimulus which repeatedly precedes the occurrence of the reinforcing stimulus will acquire discriminative control over members of the class.

(Manuscript received May 28, 1947)

References

1. Estes, W. K. Discriminative conditioning. I. A discriminative property of conditioned anticipation. *J. exp. Psychol.*, 1943, **32**, 150–155.
2. Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *J. exp. Psychol.*, 1941, **29**, 390–400.
3. Skinner, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton-Century, 1938.
4. Walker, K. C. Effects of a discriminative stimulus transferred to a previously unassociated response. *J. exp. Psychol.*, 1942, **31**, 312–321

THE EFFECTS OF MILD ANOXIA ON SIMPLE PSYCHOMOTOR AND MENTAL SKILLS *

BY ROGER W. RUSSELL

Department of Psychology, University of Pittsburgh

I. Introduction

The presence of anoxia is widespread among many varieties of abnormal behavior in living organisms. Its exact role is not clearly evident for it may be a causal factor or it may only be related to other aspects of the behavior, all being manifestations of some more basic phenomenon. McFarland (**8, 9, 10**) has conducted extensive studies of the effects of anoxia on many aspects of human behavior and of the behavior changes accompanying acclimatization to oxygen want. Bills (**3**), in studies comparing blocking in mental fatigue and anoxemia, has suggested that the former may be due, in part at least, to the reduction of the oxygen supply to the functioning tissues. Himwich and his collaborators (**5, 6, 7**) have stressed the importance of cerebral anoxia in schizophrenia, and other forms of abnormal behavior. Gellhorn (**4**) and his co-workers have studied the effects of anoxia in induced convulsive seizures and in relation to certain characteristics of the conditioned response. From these studies he has proposed that, in the case of shock therapies for mental disease, ". . . the beneficial effect is based on the development of a reduced rate of oxidation in the brain with subsequent stimulation of the sympathetic division of the autonomic nervous system" (**4**). These, and many other research reports, have indicated that oxygen changes affecting metabolism may be fundamental physiological factors in a variety of abnormal behavior patterns.

The present study [1] is one of a series originally motivated by the practical need to learn the effects of anoxia on the various skills required of the heavy-bombardment air-crew. During the course of these studies it became apparent that there occurred behavior changes to short as well as long exposures to anoxia. It is believed that the observations to be reported supplement the detailed and extensive information available on long duration anoxic effects and add to an understanding of the role of anoxia in abnormal behavior patterns.

Two hypotheses had been developed on the basis of casual observations of behavior under anoxic conditions: (1) Mild anoxia causes

[1] The data reported here were obtained by the author during a project of the AAF Aviation Psychology Program.

an immediate decrement in the performance of simple psychomotor and simple mental skills; and (2) rapid adjustment occurs in these simple skills as the time under mild anoxia increases. These hypotheses have been tested and both appear to be substantiated by the results.

Previous studies have contributed to the testing of these two hypotheses, although the samples have been limited in size and, frequently, desirable controls have not been operative. Also the majority of studies have been concerned with the effects of anoxia over periods of hours or days, rather than minutes as is the case in the present study of short-duration effects. The results [2] tend to indicate that: "Simple sensory and motor responses are not seriously impaired until the subject approaches collapse from O_2 want, and then the loss appears to be fairly sudden" (8). "The effect is more marked in mental than in physical work and the more habitual the task the less the effect" (14). There exists an ". . . unusual capacity of the organism to become adjusted to a deficit of oxygen in the inspired air, provided a long enough time is allowed for acclimatization to take place" (10). ". . . There are very definite deteriorating effects due to a reduction in oxygen pressure for unacclimatized man which may be first manifested in the neighborhood of 12,000 feet and which became marked at 15,000 to 17,000 feet . . ." (10). Most of the investigations from which these results have been culled were designed to study deterioration and adaptation over relatively long periods of time with varying degrees of anoxia. The present research deviates from them in that the time dimension over which anoxia effects have been noted was relatively short, i.e., 35 min. in all instances reported. These short-duration effects may prove to be of as great interest to the student of behavior as are the relatively long-duration effects previously studied.

II. Method and Procedure

The method of difference design used in the research is diagrammed below. Three dependent variables were investigated: (1) finger dexterity; (2) arm-hand coordination; and (3) simple addition. Finger dexterity was measured in terms of the number of pegs turned on a

Group	Initial	Intermediate	Final
Experimental	Ground altitude (Normal)	Simulated 18,000 feet altitude (Anoxic)	Ground altitude (Normal)
Control	Ground altitude (Normal)	Ground altitude (Normal)	Ground altitude (Normal)

[2] For detailed summaries of the history of anoxia and its psychological effects the reader is referred to McFarland (8) and Van Liere (15).

peg-board in a two-min. time period; arm-hand coordination, by the number of pegs moved from one position to another on a pegboard in a similar period; and, simple addition, by the number of problems completed correctly in a three-min. period. A more complete description of these performance measures will follow. The first two have been classed as simple psychomotor skills, and the third, as a simple mental skill.

In this study mild anoxia is defined in terms of the barometric pressure to which the experimental subjects were exposed and the length of time during which they were exposed, i.e., 380 mm. Hg (equivalent to 18,000 feet of altitude) and 35 min. exposure. This is equivalent to one-half the barometric pressure at sea level (760 mm. Hg), the partial pressure of oxygen under these experimental conditions being 80 mm. Hg. At no time during these experiments did the anoxic condition of even a single *S* reach the point of causing a neuro-circulatory collapse. Van Liere (15) in his monograph on anoxia warns that: "In discussing the effects of oxygen wants on physiologic processes, it is generally recognized that four factors determine the response of an organism under low oxygen tension: (1) the suddenness of the production of oxygen want; (2) the degree of severity of oxygen want; (3) the duration of oxygen want; and (4) the physical condition of the body." It will be seen in examining the procedure followed in the present experiments that all possible precautions were taken to ensure that these variables were controlled throughout each phase of the research.

Great care was taken to control all relevant variables other than the independent and dependent variables.

1. *Selection of subjects.*—A total of 244 *Ss* participated in the experiments, 67 in the finger dexterity group, 81 in the arm-hand coordination group, and 96 in the simple addition group. Each group was further divided into two sub-groups, i.e., experimental and control. All *Ss* were Air Force cadets undergoing classification of the Santa Ana (California) Army Air Base. Because of the rigid requirements for cadets, the *Ss* constituted a carefully selected and relatively homogeneous group in terms of their physical condition. All were between the ages of 18 and 27 years. They had taken at least one previous 'flight' in the high altitude chamber and were acquainted with procedures followed as well as with the conditions associated with decreased barometric pressure. No *S* was selected unless he volunteered for the experiment and no single *S* was allowed to serve in more than one experimental session.

2. *Initial equality of groups in the skills studied.*—Experimental and control groups for each skill were comparable on the basis of initial performance scores. Table I shows their comparability.

TABLE I

COMPARABILITY OF GROUPS IN INITIAL LEVELS OF PERFORMANCE

Group	Finger Dexterity		Arm-Hand Coordination		Simple Addition	
	(N = 43) Expt.	(N = 24) Control	(N = 50) Expt.	(N = 31) Control	(N = 48) Expt.	(N = 48) Control
Mean	40.06	40.10	35.91	36.78	24.12	25.12
σ_M	0.69	0.81	0.38	0.49	0.84	1.18
CR	0.04		1.40		0.69	

3. *Comparability of experimental and control situations.*—All experimentation was conducted in the high altitude chamber at the Santa Ana Army Air Base. This chamber was of the standard type used by the Army Air Forces in which high altitude conditions were simulated by the reduction of barometric pressure. The chamber was operated by a trained crew under the supervision of medical officers. The *Ss* were seated in the chamber and standardized instructions concerning the measuring of the skill under study were given. Initial trials (three in the case of finger dexterity and arm-hand coordination and one in the case of simple addition) were run in the chamber at ground level. The experimental *Ss* were taken to a simulated altitude of 18,000 feet at a rate of climb of approximately 6000 feet per min. Immediately on reaching this altitude a fourth trial was run. Seven subsequent trials at that altitude followed at five-min. intervals, the com-

pletion of each trial coinciding with the end of each interval. Hence, eight trials at 18,000 feet comprised the experimental trials obtained under anoxic conditions during a total elapsed time of 35 min. Immediately following the eighth trial (18,000 feet) descent was begun at a rate of 3000 feet per min. At ground level additional trials, corresponding in number to the initial pre-anoxic trials, were run in the chamber.

The control *Ss* were treated in exactly the same manner with the exception that they were not subjected to anoxic conditions. Following the initial trials at ground level they were told that they were to ascend to the experimental altitude. They were then taken rapidly to an altitude of 500 feet and slowly returned to ground level, arriving there after exactly the same time required by the experimental group to reach 18,000 feet. This procedure was followed in order to control for any suggestion factors which might have been associated with the flight itself and which might have influenced the results. The subterfuge appeared to serve its purpose, for check lists indicated that control *Ss* believed the flight to the experimental altitude actually had been made. Trials followed exactly the same time schedule as that of the experimental *Ss*.

4. *Controls during the experimental periods.*—Care was taken to standardize instructions and testing procedures for all experimental and control groups. The *E* took oxygen for the entire anoxic period to eliminate the possibility of errors in the administration of the trials and in the recording of the results. Oxygen equipment was also worn during the corresponding control period.

5. *Checks on physical condition of subjects.*—All *Ss* were checked by medical officers prior to entrance into the chamber. No *Ss* with head colds or other temporary disorders, which might possibly affect the experiments, were allowed to participate. No neuro-circulatory collapses occurred during the entire series of experiments.

6. *Apparatus.*—Finger-dexterity was measured by performance on a square-peg board test. The apparatus consisted of a conventional peg-board fitted with 48 square pegs with round tops. The *S* was required to remove each peg from its hole in the board, rotate it through 180° in a clockwise direction by means of the thumb and forefinger, and replace it in the hole from which it had been removed.

The apparatus for measuring arm-hand coordination consisted of a board containing 48 holes and 44 dowel-pins. The dowel-pins were arranged in such a way that one row of holes in the center of the board was vacant. The *S* was required to transfer the pins from one-half of the board to the other, according to a prearranged sequence. No rotation of pins was necessary.

The tests of simple addition included 50 problems each. A problem consisted of a row of 12 digits. Preceding each row was an underlined total which the *S* was to find by adding the digits from left to right. *Ss* were instructed to draw a short vertical line to the right of the last digit necessary to arrive at the desired total. Thus, the *S* was not required to write his answer. Seven equivalent forms of the test were available. Three forms (the first, second and third) were administered twice to make the total of 10 tests required for the experiment.

III. Results

1. *A decrement in level of performance appeared immediately after the introduction of mild anoxia.*—Figs. 1, 2 and 3 present the mean levels of performance of experimental and control groups in the three skills studied. Anoxia was introduced immediately following trial 3 in the case of finger dexterity and arm-hand coordination and trial 1 in the case of simple addition. It will be noted in all figures that a sudden decrement in performance occurred at this point, reaching an asymptotic level within two or three trials (10–15 min.).

Another method of analyzing the data on this point is to compare the performance of each individual on trials 2–14 or 2–10 with his performance on his initial trial. This method has the advantage of eliminating the variable of inter-individual differences, since each *S* serves as his own control. When such a procedure is followed, the

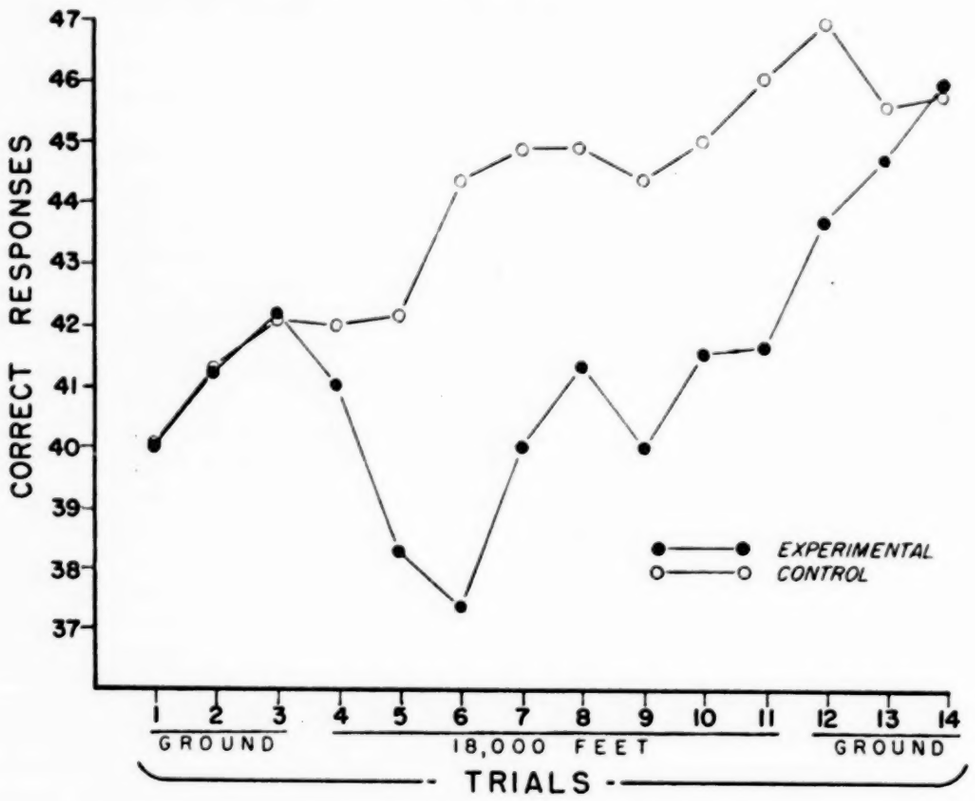

FIG. 1. Finger dexterity

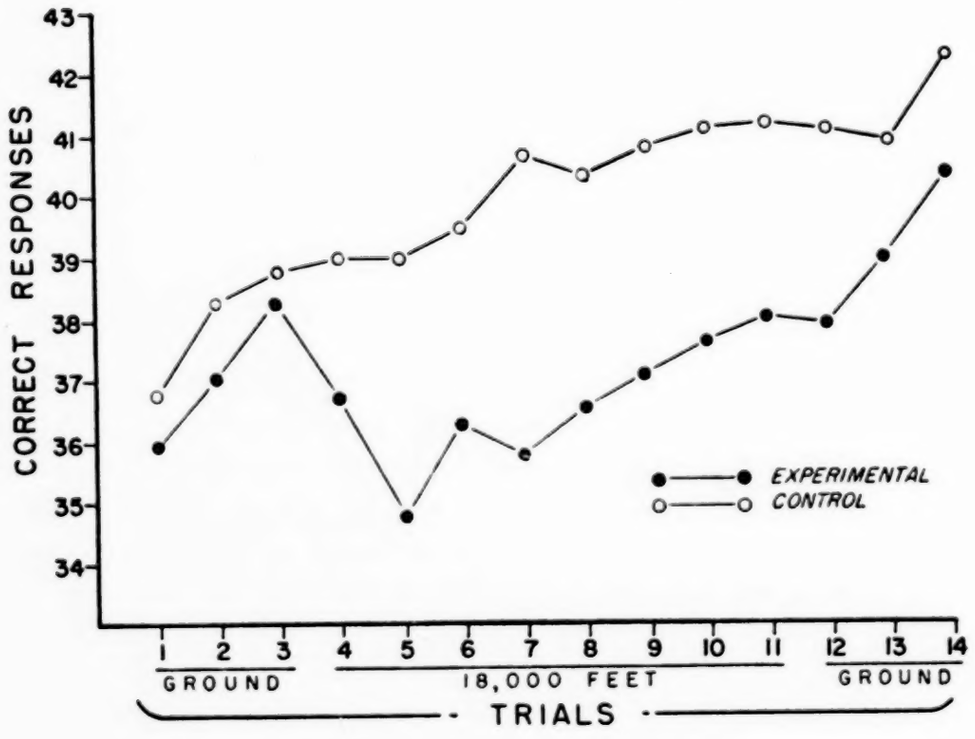

FIG. 2. Arm-hand coordination

same trend is indicated as in the previous analysis. The most significant single observation from these data is that the increasing level of performance of the control groups is paralleled in the experimental group by a decreasing level of performance during the period immediately following the introduction of anoxia.

2. *Following an initial decrement in level of performance, adjustment occurred as the time under mild anoxia increased.*—Analysis of the data presented above indicates a consistent rise in levels of performance with increasing time under anoxia following the initial decrement. This indicates an adjustment to the anoxic conditions. Theoretically, adjustment can be said to be complete when the performance

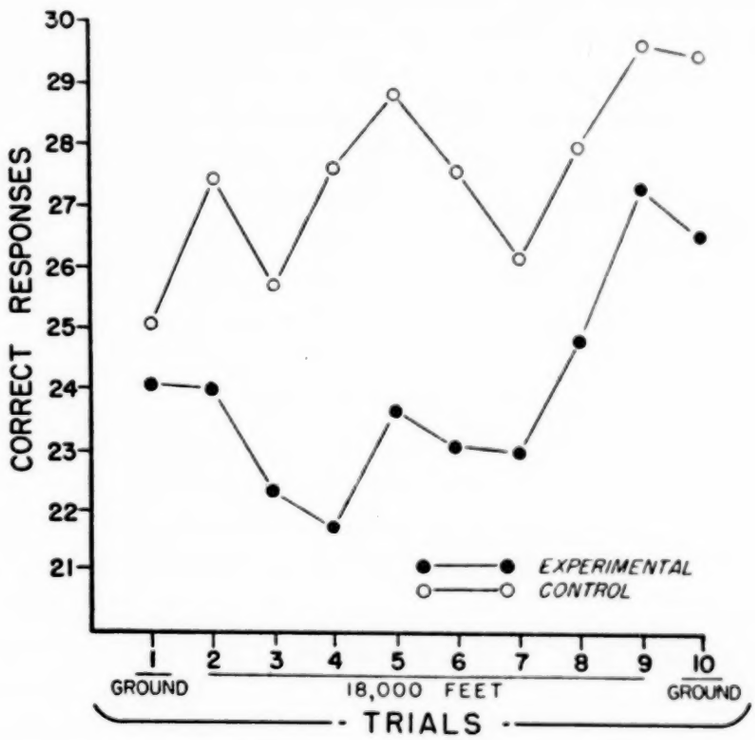

FIG. 3. Simple addition

level once again reaches the level just preceding the introduction of anoxia. It should be noted that performance actually exceeded this level, although it did not match the performance of the control during the anoxic period.

3. *In addition to the adjustment process there occurred improvement with continued practice under the anoxic conditions.*—Examination of the data previously presented indicates that more than mere adjustment occurred during the later trials under anoxic conditions. Improvement took place with practice as indicated by the fact that the final trials, even those under anoxia, were at higher levels of performance than the initial trials. These differences in performance, with the exception of finger dexterity, are statistically significant.

The slope constants for the experimental period of improvement under anoxia and for the corresponding period for the control groups were calculated. In all cases the slope constants for the experimental groups exceed those for the controls, indicating a more rapid rate of improvement with practice for the experimental Ss.

IV. Discussion and Interpretation

Within the limits of the research design the following conclusions may be drawn from these experiments.

1. *Mild anoxia produces an immediate interference in the performance of simple psychomotor and mental skills.*—The interference was observable after five min. exposure to oxygen deficiency and was exhibited by a decrease in level of performance.

This conclusion may appear to be in contradiction to the conclusion reached by McFarland (**8**) and quoted in the introduction to this report that "simple sensory and motor responses are not seriously impaired until the subject approaches collapse from O_2 want. . . ." The present impairment was observed with none of the customary signs of collapse. This impairment was statistically significant but might not be interpreted as being 'serious,' since rapid adjustment also occurred.

McFarland and Dill (**10**) report in a later study that ". . . there are very definite deteriorating effects due to a reduction in oxygen pressure for unacclimatized man which may be first manifested in the neighborhood of 12,000 feet and which become marked at 15,000 to 17,000 feet. . . ." The present results substantiate this generalization in terms of short-duration effects.

The general tenor of the present results supplements other research which has indicated significant effects of mild anoxia on other levels of behavior. For instance, Barach (**1**) and Barach and Kogan (**2**) in comparing the gross behavior of a group of students and a group of mental patients observed a ". . . lack of emotional restraint . . ." accompanying anoxia induced by a simulated altitude of 12,400 feet. Smith (**12**) and Smith and Seitz (**13**) in experimentation on speech intelligibility under conditions of anoxia conclude that, "Under standard conditions speech intelligibility is shown to decrease with increasing altitude." Curves for speech intelligibility as a function of time under anoxic conditions (10,000 feet) presented by Smith have the same general shape as those of the present study, although the time axis is extended to hours rather than min. The effect of anoxia is observable in terms of a decrement in intelligibility between the first and second tests of the experimental group.

All these observations indicate the sensitivity of the human organism to changes in oxygen supply. There are several causes for the

decrease of oxygen supply in the human organism. It is reasonable to suppose that behavior changes will be induced regardless of the specific cause of the deficiency. How extensively anoxia operates in affecting such behavior changes is a problem in need of systematic investigation.

2. *In the performance of simple psychomotor and mental skills the human organism is capable of rapid adjustment to mild anoxia.*—The exact nature of this adjustment cannot be described at the present state of knowledge. It is an empirical concept observable in terms of an improvement in performance following the initial decrement induced by mild anoxia. In reality it may represent a rapid adaptation of the organism to the oxygen want, a form of acclimatization discussed by McFarland in connection with behavior changes over a relatively long period of time. It may be a matter of increased effort on the part of the *S* as suggested by Smith (**12**) in his discussion of the 'end-spurt' observed in some of his studies. It may be the result of other changes in motivation similar to those proposed by Seitz and Keller (**11**) in their study of oxygen deprivation and conditioning in the white rat. Or the observed improvement may be accounted for by the combined action of several of these factors.

3. *This adjustment is sufficient to allow improvement with practice to occur very soon after the mild anoxia is induced.*—One of the most interesting theoretical points concerning the effects of oxygen changes on behavior is the influence exerted on the learning process. McFarland (**9**) in one of his reports of psychophysiological studies at high altitudes in the Andes stresses the importance of critical altitudes (17,500 feet and above) where significant impairment occurs in such functions as rote memory, associated memory and intraserial phenomena. On the other hand, Seitz and Keller (**11**) report experimental results on oxygen deprivation and conditioning in the white rate which ". . . would seem to indicate no discernible effect of low oxygen upon the learning process itself. . . ." Their experiments, involving the conditioned lever pressing response, were conducted at a simulated altitude of 17,500 feet. The present experiment indicates that, over short periods of time, the adjustment of the human organism is sufficient to allow improvement with practice in simple psychomotor and simple mental skills to occur very soon after the mild anoxia is induced.

Why the rate of improvement was greater for the experimental than for the control groups is puzzling. One explanation appears to lie in the possibility that the control groups were approaching their performance limit on each of the tasks, while the experimental groups, due to the initial decrement in their performances under anoxia, had farther to go to approach this same limit.

V. Summary

An investigation of the effects of mild anoxia on three simple psychomotor and mental skills has been conducted. The skills studied were finger dexterity, arm-hand coordination and simple addition.

Two hundred and forty-four Army Air Corps Cadets served as *Ss*. Each *S* participated in only one experiment. They were divided into sub-groups equal in initial level of performance on the skill in question. The experimental sub-group in each case was subjected to mild anoxia, while the control sub-group was tested always under normal oxygen conditions. The mild anoxia was induced by reducing the barometric pressure in a high-altitude chamber of the variety used throughout the war by the Army Air Forces. Time under anoxia was held constant at 35 min., since the study was concerned with the short-duration effects of the anoxic conditions.

The results justify three conclusions. (1) A decrement in level of performance appeared immediately after the introduction of mild anoxia. (2) Following an initial decrement in level of performance rapid adjustment occurred as the time under mild anoxia increased. (3) In addition to the adjustment process there occurred improvement with continued practice under the anoxic conditions.

(Manuscript received June 18, 1947)

References

1. Barach, A. L. The effect of low and high oxygen tensions on mental functioning. *J. Aviat. Med.*, 1941, **12**, 30–38.
2. Barach, A. L., & Kogan, J. Disorders of mental functioning by varying the oxygen tension of the atmosphere. *Psychosom. Med.*, 1940, **2**, 53–67.
3. Bills, A. G. Blocking in mental fatigue and anoxemia compared. *J. exp. Psychol.*, 1937, **20**, 437–452.
4. Gellhorn, E. Effects of hypoglycemia and anoxia on the central nervous system. *Arch. Neurol. Psychiat.*, 1938, **40**, 125–146.
5. Himwich, H. E., Alexander, F. A. D., & Lipetz, B. Effects of acute anoxia produced by breathing nitrogen on the course of schizophrenia. *Proc. Soc. exp. Biol. Med.*, 1938, **39**, 367–369.
6. Himwich, H. E., Bowman, K. M., Daly, C., Fazekas, J. F., Wortis, J., & Goldfarb, W. Changes in cerebral blood flow and arterio-venous oxygen difference during insulin hypoglycemia. *Proc. Soc. exp. Biol. Med.*, 1940, **45**, 468–469.
7. Himwich, H. E., & Fazekas, J. F. Factors of hypoxia in the shock therapies of schizophrenia. *Arch. Neurol. Psychiat.*, 1942, **47**, 800–807.
8. McFarland, R. A. The psychological effects of oxygen deprivation (anoxemia) on human behavior. *Arch. Psychol.*, N. Y., 1932, **145**, 1–135.
9. McFarland, R. A. Psychophysiological studies at high attitudes in the Andes. Part III. Mental and psychosomatic responses during acclimatization. *J. comp. Psychol.*, 1937, **24**, 147–188.
10. McFarland, R. A., & Dill, D. B. A comparative study of the effects of reduced oxygen pressure on man during acclimatization. *J. Aviat. Med.*, 1938, **9**, 18–38.

11. SEITZ, C. P., & KELLER, F. S. Oxygen deprivation and conditioning in the white rat. I. The effect of deprivation upon the lever pressing response. *J. Aviat. Med.*, 1940, **11**, 210–213.
12. SMITH, G. M. The effect of prolonged mild anoxia on speech intelligibility. *J. appl. Psychol.*, 1946, **30**, 255–264.
13. SMITH, G. M., & SEITZ, C. P. Speech intelligibility under various degrees of anoxia. *J. appl. Psychol.*, 1946, **30**, 182–191.
14. TANAKA, K. Experimental study on the effects of low barometric pressures and oxygen deprivation upon the efficiency of mental and physical work. *Rep. Aero. Res. Inst. Tokyo Imperial Univ.*, 1928, **3**, 128–230.
15. VAN LIERE, E. J. *Anoxia: its effects on the body.* Chicago: Univ. Chicago Press, 1942.

VISUAL FIELD ARTICULATION IN THE ABSENCE OF SPATIAL STIMULUS GRADIENTS

BY CARL R. BROWN

University of Michigan

AND

J. W. GEBHARD

Johns Hopkins University

INTRODUCTION

It has been pointed out (**11**) that intermittent, spatially uniform stimulation of the entire, or nearly entire, retina gives rise to an articulated visual field of considerable complexity. It was suggested that this phenomenal articulation be considered the general manifestation of random metabolic variations in the visual mechanism so that *patterns* were developed in excitation, notwithstanding uniform stimulation. This hypothesis will be considered in greater detail later in this paper. The observations reported here attempt a further examination of these visual 'transient' phenomena.

Purkinje (**24**), in a remarkable work, reported a large number of subjective effects produced by fixating a bright surface and waving the fingers rapidly before the eyes to produce intermittence. These so-called 'Licht-Schattenfiguren' he grouped into primary and secondary forms. The 'primary' figures, best seen with eyes open toward a bright surface, include: (a) large and small quadrilaterals (*Vierecke*) alternately light and dark, making up a checkerboard pattern (*Würfelbrett*); (b) shifting zigzag lines; (c) hexagons toward the bottom of the field with gray contours and white interiors; (d) a rose leaf pattern (*Rosettfigur*) in the center of the field. Closed eyes in direct sunlight disclose with intermittence the secondary patterns: a rectangular spiral (*Schneckenrechteck*) and a fan-like form composed of eight rays (*Achtstrahle*).

Thomsen (**27**) gave considerable attention to Purkinje's subjective figures and has attempted an explanation of the primary patterns. He suggested that the quadrilaterals may be explained by assuming that the retina is divided functionally into quadrilateral fields which somehow become visible with intermittence. The zigzag lines were held related to similar phenomena observed in flicker scotoma. The hexagons he treated, with Wolffberg (**30**), as the entoptic visibility of the hexagonal cells of the pigment epithelium. Finally, the rosette figure was considered an outgrowth of the retinal capillary arrangement.

Concerning the secondary figures, Thomsen made no positive suggestions.

Helmholtz (13) referred to Purkinje's observations with little comment, but described his own perceptions when pressure was exerted against the eyeball. Purkinje had also systematically described pressure images (*Druckphosphene*) and found them to contain many of the patterns produced by intermittence. Helmholtz, however, noted something more nearly like certain phenomena to be reported below. He saw at first a finely patterned field of diverse colors which sometimes appeared as though strewn with fine leaves or moss. This then became a pattern of bright brownish-yellow quadrangles and fine lines, and lastly developed into dark lines on a brownish-yellow background. It sometimes appeared as complex, star-shaped figures and a weaving, maze-like pattern. He also reported the presence of bright blue and red sparks in certain parts of the field.

Müller (20) studied irregular and non-specific images in the light and dark visual field that are not dependent upon external stimulation. He believed them due to influences of other physiological systems upon the visual apparatus.

Procedure and Results

The following observations were carried out in a dark room with stimulation of the left eye: the right eye was covered, except as indicated below. The authors were the two observers.

Experiment I. Differentiation of Two Phases

Preliminary observations were made with the aid of a flashed opal diffusing screen as near to the eye of *S* as possible. Illumination was provided by a projection system. An episcotister with two 90° open and two 90° closed sectors was placed in the light beam at the point of focus and was driven by a variable speed motor.

Observation of the visual field at any flash frequency from two flashes per sec. (f.p.s.) to 20 disclosed a most remarkable and beautiful display of color and form perceptions. These were roughly similar to Purkinje's and Thomsen's descriptions, but important differences are to be noted. In agreement with Helmholtz, the effects seen by the authors did not show the regularity implied by Purkinje and Thomsen. The visual field was comprised of forms in a continual state of motion and only momentarily could any particular pattern be observed. Rarely did any single pattern completely dominate the field. This involved state of perceptual flux at first appeared merely kaleidoscopic until careful observation distinguished two phases, a bright and a dark. These underwent a slow, irregular fluctuation.

Bright phase.—During the bright phase the field was dominated by certain characteristic geometrical patterns. These were sometimes in the form of windmill vanes (probably the *Achtstrahle* of Purkinje and the *Mühlenflügel* of Thomsen) and sometimes in the form of a checkerboard (*Würfelbrett*). They were alternately yellow

and blue of remarkable saturation. From time to time fleeting bands of hues arranged themselves across the field in a form suggestive of a rainbow. It was impossible to ascertain the number of distinct hues in the bands, nor could the order of hues be determined. Apparently all hues were present. They appeared in all positions of the field, sometimes appearing to originate in the center of the field, sometimes not.

The chief disagreement between the observers was with respect to the hexagons. One experienced them easily, but the other saw them only rarely and then indistinctly.

Dark phase.—This phase was much simpler and consisted of a writhing mass of violet flecks against a yellow-green background. The areas of violet and yellow-green were nearly equal—perhaps the violet area was larger—yet the yellow-green always appeared to be the ground. The violet flecks were observed as being irregularly angular in shape, and the patches which they made up seemed to become coarser as the flicker frequency increased. Their constant motion was rapid and curvilinear.

Experiment II. Effect of Contour

Upon the assumption that there is a direct relationship between heterogeneity in the stimulus field and segregation, an effort was made to provide a physically homogeneous stimulus condition in order to study the phenomenon in its simplest form. An attempt was made (1) to produce a field having no external limits other than those imposed by the margins of the orbit,[1] and (2) to achieve a field free from texture and other inhomogeneities.

A short piece of thin-walled brass tubing three in. in diameter was fitted snugly over the left eye by shaping to fit the irregularities of the face. A layer of dentist's wax was placed around the shaped edge and, after warming to soften, was molded exactly to required form by moderate pressure against the parts of the face surrounding the eye. The other end of the tube was recessed to hold a flashed opal diffusing screen placed as close to the cornea as possible. This device was clamped to *S*'s face and the chamber between the screen and the cornea was filled with clear mineral oil. Since the refractive index of the mineral oil (1.48) is greater than that of the cornea (1.376), the effect is to transform the cornea into a weak negative lens. Since the cornea normally contributes about 43 diopters of the total 59 diopters of the eye's refracting power it was hoped that the oil chamber would virtually eliminate both boundary and texture of the field. The superior and inferior margins were entirely invisible, but in the temporal and nasal regions, where the visual angle is greatest, the edges of the field were still visible. The external surface of the flashed opal screen was illuminated by a powerful beam of parallel rays and it is believed that this method eliminated intensity gradients within the field of vision. The light was interrupted by a variable speed episcotister as before and no light was permitted to enter the right eye.

---

[1] Previous attempts to achieve a total stimulus field have employed rather cumbersome methods. Aubert (2) had *S* sit close to a small illuminated screen and depended upon convergence to make the screen area fill the entire visual field. Dunlap employed a large, internally illuminated sphere in which the head was placed, and Garten (10) designed a cylindrical field illuminated from behind. Most recently, Engel (7) filled the entire visual field with a large wall surface, but Metzger (19), who used the same method, was not able to escape the problem of texture at high illuminations. It was thought that a more compact 'total field' could be constructed, and because the attempt reported here was not completely successful, research is now being undertaken with a more promising method.

Under these better controlled conditions the perceptual phenomena were not different from those observed in Experiment I. The two-phase effect was still the most conspicuous aspect of the observations. Apparently, the extent to which field boundary and texture were reduced by this procedure had no appreciable effect on the perceptual segregation.

Experiment III. Binocular Rivalry

In experiments I and II the left eye was stimulated and the right eye was dark. The slow alternation of the two phases suggested binocular rivalry so strongly, in spite of the contradiction offered by the supposed uniocular stimulation, that steps were taken to obtain additional evidence. Additional apparatus in the form of a plaster of Paris cast to fit the upper half of *S*'s face was made. The brass chamber described above was set into the cast over the left eye and a smaller tube fitted with a light-proof screw cap was set into position over the right eye. A small hole was drilled into the right hand tube in such a position that light could be projected into the right eye at such an angle as to produce a weak, diffuse illumination of the nasal periphery. Employing the well-known procedure in binocular rivalry of introducing a change in one monocular field to make it the dominant one, the right eye was briefly illuminated at irregular intervals and the relationship to the phases was noted. It was observed that whenever the bright phase was dominant, stimulation of the right (dark) eye brought instant dominance of the dark violet and yellow-green pattern. It seems evident, therefore, that the violet and yellow-green pattern is a function of the externally dark eye and that the phenomena of the bright phase occur in that portion of the optic sector served by the directly stimulated eye. Is the 'externally dark' eye nevertheless stimulated? We consider this question in a later section.

Experiment IV. Effect of Intermittence, Composition and Intensity of Illumination

It has been stated that perceptual segregation occurs under steady homogeneous stimulation, but is greatly enhanced by intermittent illumination. As already noted, with steady light fleeting bands of colors are seen in the visual field. Described more accurately, they consist of multizone bands of approximately 5°–10° total width, which appear now in one meridian, now in another. All spectral hues and purple are represented in low saturation and indeterminate order. They appear strikingly rainbow-like, in spite of the peculiar uncertainty of the order of the various hues. However, the facts of the

appearance of these 'rainbows' in different meridians at different times, and of their presence when filtered light is used, seems to be sufficient evidence that they are not due to physical artifacts. The angular violet flecks against the yellow-green background are visible also. With slow intermittence, two f.p.s., the phenomena of both the light and dark phases become clearer and more numerous. They remain, however, essentially the same. As the flash frequency is increased up to 20 f.p.s., the violet and yellow-green pattern of the dark phase appears to become coarser.

The effects so far described were obtained with unfiltered, almost white tungsten light providing a field brightness at the flashed opal diffusing surface of approximately 200 millilamberts. A series of observations at various flash frequencies was also undertaken with red and green light.[2] The red field had a brightness of 1.5 millilamberts and the green of 12 millilamberts, approximately. With both the field articulation was similar to that observed with white light. The familiar field desaturation phenomenon was observed and against the resulting gray the vari-colored bands and the violet and yellow-green pattern were clearly seen. Evidently field articulation does not depend upon the composition of the illuminant.

Even more remarkable is the fact that chromatic perceptions persist when the intensity of the stimulus light is below the cone threshold. With red illumination of the limited wave-length range already referred to, and intensity below the chromatic threshold, the two-phase effect under intermittence was present in an essentially unaltered form. The suggestion was made in Experiment III that the violet and yellow-green pattern of the dark phase is related to activity in the externally dark eye. Corresponding to this is the further observation that this pattern is continuously visible in total darkness to both authors for periods up to an hour or more. A close examination of the perceptions of the so-called 'light chaos' revealed its presence in a clear and unvarying form throughout the course of dark adaptation and after.

Experiment V. Relation of Dark Phase to Bioluminescence

The ubiquitous violet or blue images observable in many visual phenomena have attracted the attention of numerous observers. Helmholtz (13) mentioned a 'watered silk' effect in intermittent illumination. Ives (14) found, in her investigation of the flight of colors, that the intrusion of a violet phase interfered with the chromatic order of the after-images. The color sequences would have been almost completely regular and reproducible

[2] The chromatic illumination was the same used in a previous investigation (11). The red contained no wave lengths shorter than 650 mμ and the green had a spectral centroid of approximately 560 mμ for ordinary incandescent illumination.

under the same conditions, if the violet phase could have been omitted. Young (31) reported the presence of purple or rose-hued clouds on the surface of a rotating mirror in his differential color mixer. He tentatively suggested that this might be subjectively visible rhodopsin. Bidwell (3), in experiments on the 'ghost,' attributed the violet coloration of this after-image to the postulated V fibers—processes—of the retina. Most recently, Friedman (9) has called attention to the prevalence of blue in entoptic phenomena.

Judd (15), to account for the three stages of the Purkinje after-image encountered in his work on that problem, suggested a hypothesis of secondary stimulation referred to as *bioluminescence*. This hypothesis originated in an attempt by Druault (5), and later by Ladd-Franklin (18) independently, to explain the phenomenon of the blue arcs of the retina, an effect which may well be related to certain aspects of the violet and yellow-green pattern. According to this view, it is assumed that some retinal structure gives off, after excitation, visible light of a whitish-violet character.[3] It is important to note in this connection observations by Gertz (12) and Ladd-Franklin (18), who claim that the blue arcs are followed by a complementary after-image.[4] Since the arcs are a reddish-blue, their after-image would be a yellowish-green. These hues, insofar as can be judged, correspond to those of our violet and yellow-green pattern. Upon the assumption that the violet aspect might be due to a retinal emission effect and that the yellow-green was the residual image of it, an attempt was made to obtain direct evidence of any radiations emitted by the unstimulated retina, the apparent seat of the phenomena of the dark phase.

Direct examination of the right, dark (so far as illumination through the pupil is concerned) retina by means of a carefully shielded ophthalmoscope, while the left retina was being stimulated by intermittent light, failed to disclose anything. Even with prolonged dark adaptation, the examiner was unable to see what was plainly visible to *S*.[5] Druault (5) had anticipated this difficulty in suggesting that since the light emitted was of short wave length it would probably be absorbed by the ocular media. Here, since the eye of the examiner was also involved, the increased absorption of the radiation—if any—might doom the venture at the outset.

It was then thought that some photographic indication of retinal emission might be obtained by fitting the right eye-piece of the plaster

[3] The *bioluminescence* hypothesis, supported by Druault (5), Ladd-Franklin (18), and Judd (15), is an alternative to a view of secondary electrical stimulation held by Gertz (12), Troland (28), Amberson (1), Newhall (21) and others. The theoretical problem of the blue arcs is reviewed by Friedman (8).

[4] Newhall (21) questions the sufficiency of the evidence here, since Gertz and Ladd-Franklin themselves are the only persons who have ever reported the residual image. None of their *S*s did, nor did any of Newhall's.

[5] Ellis (6), observing directly a frog's retina, failed to see any visible light accompanying stimulation.

cast described in Experiment III with an orthochromatic emulsion.[6] In anticipation of light leakage from one eye to the other through the head, the left eye tube was equipped with a red filter transmitting no appreciable radiation shorter than 650 mμ. Since orthochromatic emulsions are relatively insensitive to red light, this was thought to be something of a safeguard.

Exposures up to 30 min., while the left eye was being intermittently stimulated, were made. The 30 min. exposures produced some faint fogging of the film, but the experimental conditions do not permit attributing this to light emitted from the retina. Consideration must be given to the possibility of leakage of light from one orbit to the other. Skin and muscle tissues are by no means opaque to light and the contents of the orbit, while dense, cannot be expected to absorb all radiation. The medial wall of the orbit is to a large extent comprised of an extraordinarily thin bone, such that its name is derived from that fact (lamina papyracea). This bone lies between the orbit and the ethmoidal cells, which in turn separate the orbit from the upper portion of the nasal fossa. The ethmoidal cells, usually 8 or 10 in number, are separated from each other by thin, bony partitions.[7] This arrangement of delicate bones and air cells which open upon one another is not an ideal absorption medium and the effectiveness of the long exposure time to produce fogging of the negative may well be due to light leakage. The adequacy of the control—use of red light—is something less than ideal, since orthochromatic negative materials are not blind to the red light used, but only relatively insensitive to it for short exposures. If appreciable interocular light leakage is a fact, it would have important bearing on the apparent factor of binocular rivalry described in Experiment III, and the bioluminescence hypothesis becomes extremely dubious.

Discussion

The conditions which give rise to perceptually articulated visual fields are undoubtedly of many levels of complexity ranging from total darkness to the extraordinarily complex stimulations which are met in every-day seeing. In total darkness, as is well known, articulation is not lacking. The so-called light-chaos, or intrinsic light, is visible as a series of allegedly achromatic waves and currents. According to Helmholtz, the self-light is always whitish, but Friedman (9) has reported the presence of blue bands, and the authors have

[6] Ladd-Franklin noted, as corroborative of her view, that living tissue had been photographed by its own light by Nodon (22). Davis (4), in a criticism of Ladd-Franklin's explanation of the blue arcs, cites a review by Keenan (17) bearing on the photographic evidence. According to the latter, Nodon's 'radiations' appear to be due to the evolution of hydrogen peroxide.

[7] See Whitnall (29); especially pp. 48–55 and p. 63.

found the phenomena of the dark field to be characterized in their perceptions by the motile violet and yellow-green pattern. Next in terms of stimulus complexity might be a stimulus field completely without contour and illuminated homogeneously with steady light of an intensity below the cone threshold. Nearly on the same level with this condition is that in which the homogeneous field is illuminated with steady monochromatic light, such as red, so chosen as to affect only one of the photochromatic receptor processes. A further step is the use of light whose composition stimulates more than one of the photochromatic processes. With these last three conditions faint bands of hues appear in all meridians to give the non-spectral rainbow effect. The violet and yellow-green pattern in the unstimulated eye is as usual. Use of white light does not appreciably alter the perceptions. When the illumination is made intermittent, the phenomena just mentioned become clearer and several new forms appear. A spoke-like figure (the *Achtstrahle*) of alternate yellow and blue is seen, which occasionally gives way to a checkerboard (*Würfelbrett*) effect. Special patterns of intermittence such as are used to produce the Fechner-Benham phenomena and stationary stimulus arrangements such as the Luckiesh-Moss line patterns also must be included somewhere in the catalog.

Articulation, then, not known at present to be specifically dependent upon retinal structure, occurs under the very simplest stimulus conditions that have been studied. While the results reported here are not in complete agreement with Purkinje (**24**), Thomsen (**27**), and others, all investigations indicate some regularity and constancy in the basic patterns observed. For Purkinje and Thomsen these are chiefly the checkerboard and windmill effects, although certain others are mentioned. For the conditions observed here the 'windmill' pattern with indistinctly contoured blue masses between the yellow arms, or bands of vari-colored hues and the mottled violet and yellow-green pattern are most characteristic (Experiment I). Of the two, the violet and yellow-green pattern appears to be the more general. It is present in total darkness and is demonstrated to be the phenomenon of the externally dark eye under conditions where only the other eye is directly stimulated. The rainbow colors require that the retina be illuminated, although they may develop below the cone threshold, with and without intermittence, and with chromatic illumination as well as white light.

We thus have to try to explain at least two well differentiated kinds of field articulation, to which there are no known corresponding stimulus patterns. This problem is complicated further by the fact that one kind of field articulation seems to depend in some obscure way upon an eye which is completely shielded from external light.

Taking the latter point first, we have to consider the following hypotheses: (a) the violet and yellow-green, or dark-eye (Experiment III) pattern is of exclusively central origin. But if so, the mechanism responsible for it must be at the monocular level of function, and in any case no clue to the explanation of the pattern and its hues appears. (b) The dark-eye pattern is of peripheral origin through some process of secondary excitation which could occur only if the efferent fibres of the optic nerve are excited. (c) The retina of the dark eye is actually stimulated by light leaking into it through the medial walls of the orbits, and other intervening structures from the other, directly illuminated eye.[8]

With regard to (b), it may be well to consider whatever similarity that may exist between the violet and yellow-green pattern and the phenomenon of the blue arcs. According to Ladd-Franklin (**18**), the hue of the arcs is reddish-blue. This is denied by Snell (**26**), who claims they are blue, but if the arcs were reddish-blue their hue might correspond fairly well to what we call violet in the violet and yellow-green pattern. Further, Ladd-Franklin and Gertz assert that they have seen a yellow-green after-image of the arcs. This might then correspond to the yellow-green ground against which the violet mottling is seen. Arguing against any such parallelism are the facts: (1) that the arcs are of short duration and occur only in the nasal retina between the macula and the disc, whereas the violet and yellow-green pattern is continuously visible in all portions of the visual field; and (2) the violet and yellow-green pattern we are considering is clearly not a phenomenon of temporal induction (after-image) but one of spatial induction, if an 'induction' phenomenon at all. Nevertheless, the violet and yellow-green pattern may possibly be a less striking and more general manifestation of the same phenomenon. The most widely accepted explanation of the blue arcs is electrical. According to this hypothesis, the arcs are due to secondary excitation of adjacent retinal processes by the nerve currents resulting from the initial stimulus. The violet and yellow-green pattern may be due to analogous local electrical effects. But if this is so, the question of why such secondary excitation produces the hues violet and yellow-green remains to be answered. It is suggested by Friedman (**8**) that the blue of the arcs is due to the lower threshold of excitability of the blue process. But this does not help to account for the yellow-green. It will be remembered that Newhall (**21**) questioned the existence of a residual image following the blue arcs. Ladd-Franklin used this image in her argument for bioluminescence. Contending that phe-

[8] This possibility is not excluded by our Experiment V which was specially designed to test the bioluminescence hypothesis. The incomplete opacity of the skull bones and soft tissues in the vicinity of the orbits is easily demonstrated by the simple experiment of shutting oneself in a dark room and inserting the luminous end of an ordinary flashlight into one's mouth.

nomena produced by electrical stimulation of the visual system are not accompanied by residual images and that light only produces negative after-effects in the receptor cells, she insisted that visible light must be generated in a nerve fiber during the passage of the nerve impulse.

This second explanation of the violet and yellow-green pattern in terms of bioluminescence would be acceptable if visible radiations generated in the retina could be experimentally demonstrated. Such, however, has not been the case.

With regard to the electrical theory of secondary stimulation of the photochemical processes, proposed to account for the blue arcs, Friedman (**8**) properly points out that this would imply a backward flow of excitation from the ganglion cells to the rods and cones. Since the blue arc effect exhibits no latency, according to Newhall (**21**), the backward flow must occur simultaneously with the forward excitation. Some doubt may be raised as to whether this can occur. This difficulty might possibly be overcome with better provision of a mechanism for the violet and yellow-green pattern and the rainbow effect, by invoking the services of the centrifugal fibers of the optic nerve. The existence of efferent fibers in the optic nerves has long been known, but no definite function has ever been ascribed them, according to Cajal (**25**). The precise connections of these fibers is in doubt, but if it is accepted that they associate the cortex with some species of amacrine cell, *viz.*, the i-neurone—Polyak (**23**)—they would provide an arc of reflex stimulation of the retinal mechanism. However, such a mechanism would not differentiate between the blue arcs and violet and yellow-green pattern unless many additional assumptions were made as to distribution of efferent excitation in the retina and the mode of response of the photochemical retinal processes.

The most probable choice then seems to lie between hypotheses (a) and (c), and the latter is clearly the more conservative one. It is expected that further investigation may yield additional, possibly decisive, evidence on this point.

There remains what is probably the major problem: Why does field articulation occur at all under the conditions of these experiments, and why is the number of different patterns apparently limited when one might have expected limitless variety? Reference has already been made to the possibility of random metabolic changes in the cells of the visual apparatus, with consequent variations in the excitation pattern notwithstanding uniformity of illumination of the retina, but nothing will be gained by an uncritical introduction of such a concept. Randomness could be either temporal or spatial, or both. It may be said at once that no evidence is available that would make a discussion of temporal randomness in excitation profitable.

But it might be worth while to consider the possibility of small metabolic fluctuations (functionally manifest as sensitivity fluctuation) occurring in the photosensitive cells, or in the bipolar or ganglion cells, or the horizontal and amacrine cells, or in all of these, in a spatial distribution that is (1) completely random or (2) random in small scale but patterned in large scale. Further, notwithstanding uniform brightness of the field of vision, illumination of the photosensitive cells might be non-uniform because of optical inhomogeneities of cornea, lens, humours, and the non-photosensitive strata of the retina itself. The blood vessel network of the retina requires special consideration in this connection, since, as is well known, its shadow is visible under certain conditions.

If condition (1) is the fact, the phenomenal articulations of the visual fields reported in this paper must be interpreted as revealing the existence of some central mechanism of field articulation, intrinsically biased in favor of patterns of particular kinds, or at least of particular degrees of coarseness. This, because the degree of coarseness of the phenomenal patterns is indubitably of an order of magnitude far greater than the minimum that would be expected on the basis of the facts of visual acuity.

If condition (2) is the fact, then we must find the origin of the pattern. The retinal blood vessel network might be responsible for a pattern of metabolic fluctuations, even though local adaptation (average sensitivity) compensates completely for the shadow of the blood vessels. A pattern of field articulation having this bias should at least resemble the blood vessel pattern as it is revealed by its shadow under certain other conditions, but such is not the fact. The difficulty of explaining the configurations actually observed on the basis of optical inhomogeneities in eye structure is great enough because of lack of similarity between visual pattern and known structural features, but is made even greater by the fact that there are at least two major, and very different, classes of visual patterns.

Our interpretation, admittedly tentative, therefore takes the following form: Visual field articulation in patterns of comparatively limited variety can occur in the absence of peripheral stimulus gradients (spatial), and is favored by intermittent stimulation. (The reason for this fact is still a mystery, but one is reminded of the revival of after-images by intermittent stimulation.) These visual patterns are manifestations of intrinsic bias or anisotropy of the visual mechanism, more probably physiological than physical. The bias or biases in question are recessive in ordinary vision when relatively large stimulus gradients are imposed peripherally, but become dominant when peripheral stimulation is undifferentiated.

Summary

1. When the illumination of one monocular field of vision is spatially uniform over virtually its entire extent, but temporally intermittent at a frequency below critical flicker frequency, while the other monocular field is dark, phenomenal chromatic patterns are produced.

2. Two major patterns are distinguishable, although there are several others of a relatively transient and secondary nature.

3. Of the two major patterns, one is a radiating or 'windmill' pattern of yellow and blue of relatively high brightness; the other, much dimmer, an irregular mosaic of violet and yellow-green. These two patterns appear alternately in a rhythm strongly suggestive of binocular rivalry.

4. Evidence is given that the alternation actually is due to binocular rivalry, that the yellow-blue pattern depends on the directly illuminated eye, and that the violet and yellow-green pattern depends on the externally dark eye which is stimulated by light leaking into its orbit through incompletely opaque intervening tissue from the directly illuminated eye.

5. Evidence is presented against the hypothesis of intraocular bioluminescence.

6. A tentative hypothesis of central configurational bias is suggested.

(Manuscript received for immediate publication October 13, 1947)

References

1. Amberson, W. R. Secondary excitation in the retina. *Amer. J. Physiol.*, 1924, **69**, 354–370.
2. Aubert, H. *Physiologie der Netzhaut.* Breslau: Morgenstern, 1865, 47–49.
3. Bidwell, S. On recurrent images following visual impressions. *Proc. Roy. Soc. Lond.*, 1894, 56, 132–145; On negative after-images and their relation to certain other visual phenomena. *Proc. Roy. Soc. Lond.*, 1901, **68**, 262–268.
4. Davis, H. Theory of 'visible radiation' from an excited nerve fiber. *Science*, 1928, **67**, 69.
5. Druault, A. Sur un phénomène entoptique en rapport avec le trajet des fibres optiques dans la rétine (phénomène des arcs). *J. de Physiol. Pathol. Générale*, 1914, **16**, 649–654.
6. Ellis, F. W. The nature of the stimulus which excites the blue arcs of the retina and related phenomena. *Amer. J. Physiol.*, 1928, **84**, 485 ff.
7. Engel, W. Optische Untersuchungen am Ganzfeld. I. *Psychol. Forsch.*, 1929–30, **13**, 1–5.
8. Friedman, B. The blue arcs of the retina. *Arch. Ophth.*, 1931, **6**, 663–674.
9. Friedman, B. Observations on entoptic phenomena. *Arch. Ophth.*, 1942, **28**, 285–312.
10. Garten, S. Über die Wahrnehmung von Intensitätsveränderungen bei möglichst gleichmässiger Beleuchtung des ganzen Gesichtsfeldes. *Pflügers Arch.*, 1907, **118**, 233–246.
11. Gebhard, J. W. Chromatic phenomena produced by intermittent stimulation of the retina. *J. exp. Psychol.*, 1943, **33**, 387–406.
12. Gertz, H. Über autoptische Wahrnehmung der Sehtätigkeit der Netzhaut. *Skan. Arch. Physiol.*, 1907, **19**, 381.
13. Helmholtz, H. *Handbook of physiological optics.* Optical Soc. Amer., 1924.
14. Ives, M. Flight of colors, Thesis, University of Mich., 1936.
15. Judd, D. B. A quantitative investigation of the Purkinje after-image. *Amer. J. Psychol.*, 1927, **38**, 507–533.

16. JUDD, D. B. Least retinal illumination by spectral light required to evoke the blue arcs of the retina. *Bur. Stand. J. Res., Wash.*, 1929, **2**, 441–451.
17. KEENAN, G. L. Substances which affect photographic plates in the dark. *Chem. Rev.*, 1926, **3**, 95–111.
18. LADD-FRANKLIN, C. Visible radiation from excited nerve fiber: the reddish blue arcs and the reddish blue glow of the retina. *Science*, 1927, **66**, 239–241.
19. METZGER, W. Optische Untersuchungen am Ganzfeld, II. *Psychol. Forsch.*, 1929–30, **13**, 6–29.
20. MÜLLER, J. *Über die phantastischen Gesichtserscheinungen.* Coblenz, 1826.
21. NEWHALL, S. M. The constancy of the blue arc phenomenon. *J. opt. Soc. Amer.*, 1937, **27**, 165–176.
22. NODON, A. Recherches sur la radioactivité des cellules vivantes. *C. R.*, Acad. Sci., Paris, 1924, **178**, 1101.
23. POLYAK, S. L. *The retina.* Chicago: Univ. Chicago Press, 1941.
24. PURKINJE, J. E. *Beobachtungen u. Versuche zur Physiologie der Sinne. Beiträge zur Kenntnis des Sehens in subjectiver Hinsicht.* Bd. I, Prag, 1823.
25. RAMÓN Y CAJAL, S. *Histologie du système nerveux.* Paris: Maloine, 1911, II, 366–367.
26. SNELL, P. A. The entoptic phenomena of the blue arcs. *Arch. Ophth.*, 1929, **1**, 475.
27. THOMSEN, E. Über Johannes Evangelista Purkinje und seine Werke. *Skand. Arch. Physiol.*, 1919, **37**, 1–116.
28. TROLAND, L. T. An entoptic phenomenon demonstrating the optic impulse. *Physiol. Bull.*, 1920, **17**, 55.
29. WHITNALL, S. E. *The anatomy of the human orbit and accessory organs of vision.* London: Oxford Univ. Press, 1932.
30. WOLFFBERG, L. Die entoptische Wahrnehmung der Fovea centralis und ihre Zapfenmosaik. *Arch. Augenheilk.*, 1886, **16**, 1–13.
31. YOUNG, P. T. A differential color mixer with stationary discs. *J. exp. Psychol.*, 1926, **6**, 323–343.

DISCUSSION

CONCERNING THE EFFECT OF SHOCK FOR RIGHT RESPONSES IN VISUAL DISCRIMINATION LEARNING

BY KARL F. MUENZINGER

The University of Colorado

In a recent article in this JOURNAL (3) Wischner reported some findings on the effects of shock for right and wrong responses on visual discrimination learning in a non-correction situation. Referring to certain similar experiments in a correction situation reported by myself and others he concludes that his results "are seemingly in conflict with the findings of other studies that shock administered anywhere after choice produces a general accelerating effect" (3, p. 283). I shall try to show that a proper re-statement of Wischner's findings will indicate agreement rather than conflict between his and our results.

The major differences between our experiments were as follows: We used a T-shaped discrimination box (1, p. 268) in which the stimulus cues were facing each other across the choice point, while in Wischner's apparatus (3, p. 275) the cues were mounted in the same plane. While both of us considered an animal's choice of the alley with the positive cue leading to food one 'correct trial,' we differed radically with respect to the method followed and the terminology used in case the alley with the negative cue was chosen. Our method allowed an animal which had entered the alley with the negative cue to correct its error, enter the opposite alley with the positive cue and find food in the goal box. This is what we called one 'error' or one 'wrong trial,' no matter how often the animal entered the wrong alley and retraced it before he found food. Wischner's method did not allow for retracing in the wrong alley, but made the animal go forward to the unbaited goal box on that side and from there directly to the starting point. Each entry into the wrong alley was counted by him as one 'trial' and one 'error.' In each one of our 'trials' the animal obtained reinforcement; in one of Wischner's 'trials' it did not necessarily do so, but only in the 'trials' in which the correct alley was chosen. Our method of description permitted the counting of only one single error between two reinforcements; according to Wischner's method a number of errors could be counted between two reinforcements, namely, as often as the animal chose the 'wrong' alley.

Among other differences should be mentioned the quality of the shock: Wischner used continuous D.C., while we used D.C. impulses; and the size of the groups: our groups contained 25 animals and Wischner's, 10.

Roughly speaking, we found that the shock-right group was significantly better than the no-shock group, although poorer than the shock-wrong group

as measured both by what we called trials and errors. In contrast, Wischner found that his shock-right group was poorer, but not significantly so, than the no-shock group as measured by what he calls trials and errors. It is on this array of facts that he rests his claim that his results 'are seemingly in conflict' with ours.

Any judgment of this sort should be based, as far as possible, on comparable aspects of the two experimental situations. Wischner's use of a terminology similar to ours, i.e., 'trials' and 'errors,' does not in itself produce the necessary prerequisite for comparability, as long as these terms denote different events in the two situations. I propose, therefore, to make the two statements of results comparable by reference to the number of reinforcements required for each group to master the problem.

Wischner states his results as follows (3, p. 277):

	Shock-right	Shock-wrong	No-shock	
Mean trials	159.0	104.0	152.0	(a)
Mean errors	71.3	24.2	54.6	(b)

If we subtract the errors from the trials, as Wischner did not do, we may re-state his results in terms of mean reinforcements: shock-right, 87.7; shock-wrong, 79.8; no-shock, 97.4 (c).

The following are the results in one of our experiments (2, p. 181) in terms of what we called trials per rat, or mean reinforcements: shock-right, 76.0; shock-wrong, 54.8; no-shock, 127.2 (d). (Since Wischner's values include the criterion of learning, we here added the criterion to our values also.)

The order of the groups in (a) and (b) is not the same as in (d). This is the conflict of findings to which Wischner refers. However, the order of the groups is the same in (c) and (d). That is, the conflict disappears when the results are stated in comparable terms. We may therefore claim that, *when measured in terms of reinforcements, Wischner's results tend to confirm ours.* I have left out of consideration the question of the significance of differences, which can only be guessed at. Since the difference between Wischener's shock-wrong and no-shock groups does not seem to be significant, one can hardly expect his shock-right group to differ significantly from the no-shock group.

Without going into a detailed consideration of errors also I merely wish to point out that if we had reported each single entry into the wrong alley as one error our results in terms of 'errors' as defined by him would then also have been approximated by Wischner's. This omission on our part must certainly be regarded as a serious defect in reporting the results of our experiments.

Wischner denies having found an accelerating effect of shock-right, as we have claimed, and yet he repeatedly points out "the rapid decline in the latter part of the (error) curve when the (shock-right) animals finally begin to learn" (3, pp. 279, 283).

In conclusion I wish to say that, aside from the incomplete statement of his results and the denial of an accelerating effect which was nevertheless observed, I regard Wischner's well-planned experiment as a very necessary continuation of the problem of the shock-right effect. The non-correction

situation brings out sharply the specific effects of shock as pointed out by Wischner in his discussion.

References

1. Muenzinger, K. F. Motivation in learning. I. Electric shock for correct response in the visual discrimination habit. *J. comp. Psychol.*, 1934, **17**, 267–277.
2. Muenzinger, K. F., Bernstone, A. H., & Richards, L. Motivation in learning. VIII. Equal amounts of electric shock for right and wrong responses in a visual discrimination habit. *J. comp. Psychol.*, 1938, **26**, 177–186.
3. Wischner, G. J. The effect of punishment on discrimination learning in a non-correction situation. *J. exp. Psychol.*, 1947, **37**, 271–284.

A REPLY TO DR. MUENZINGER ON THE EFFECT OF PUNISHMENT ON DISCRIMINATION LEARNING IN A NON-CORRECTION SITUATION

BY GEORGE J. WISCHNER

University of Missouri

In his paper Dr. Muenzinger (1) has attempted to show that when the results, reported in a recent study by the present writer (3), are reanalyzed, they are *not* seemingly in conflict with the findings of other studies that shock administered anywhere after choice produces a general accelerating effect. In place of 'trials' and 'errors' Muenzinger proposes reinforcements as the more comparable measure of learning and concludes "*that, when measured in terms of reinforcements, Wischner's results tend to confirm ours.*"

It is not too clear to us how one can ignore the errors an animal makes before attaining the criterion of mastery or how one can disregard the number of trials necessary to achieve a given number of successes or reinforcements. However, it appears to us that even when stated in terms of reinforcements our data support our original conclusion, and are seemingly opposed to the generalization that shock anywhere after choice accelerates learning.

The fact that the order of scores in (c) is the same as that in (d) in Muenzinger's reinforcement analysis does not seem to us to bear on the basic issue. The point to be emphasized is that in terms of all measures—trials, errors, reinforcements—the differences between our shock-right group and no-shock group are not statistically significant although the differences between each of these groups on the one hand and the shock-wrong group on the other are statistically significant. These results are to be contrasted with those reported by Muenzinger et al. (2) which show a significant difference between the shock-right and no-shock groups in favor of the former.

In our original paper it was noted that in the Muenzinger studies results have been presented in terms of errors for the first 100 trials and that a comparable analysis of our data was not presented since it showed the same general picture as the complete learning records. Below are presented the results for each of our groups for the first 100 trials in terms of errors and

reinforcements. (The fact that the first two columns do not total 100 trials is due to the few NR—no reaction—trials obtained with some of the animals in the shock groups.)

	Shock-right	Shock-wrong	No-shock
Mean errors	55.3	23.9	44.5
Mean reinforcements	42.4	75.4	55.5

These data show clearly that in terms of the reinforcement measure the shock-right group is again poorer than even the no-shock group. The difference between the no-shock and shock-right group approaches significance, the *t*-value being 2.5. The difference between the shock-wrong group and each of the other two groups is highly significant, *t*-values of 6.0 and 4.9 obtaining for the comparisons with the shock-right and no-shock groups respectively.

These results again seem to us to be contrary to the findings of other studies that shock administered anywhere after choice produces a general accelerating effect.

References

1. Muenzinger, K. F. Concerning the effect of shock for right responses in visual discrimination learning. *J. exp. Psychol.*, 1948, **38**, 201–203.
2. Muenzinger, K. F., Bernstone, A. H., & Richards, L. Motivation in learning. VIII. Equal amounts of electric shock for right and wrong responses in a visual discrimination habit. *J. comp. Psychol.*, 1938, **26**, 177–186.
3. Wischner, G. J. The effect of punishment on discrimination learning in a non-correction situation. *J. exp. Psychol.*, 1947, **37**, 271–284.